

**НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ БІОРЕСУРСІВ І
ПРИРОДОКОРИСТУВАННЯ УКРАЇНИ**

Кваліфікаційна наукова праця
на правах рукопису

КАЧМАРСЬКИЙ ОЛЕКСА ІГОРОВИЧ

УДК 004.8:631.8:519.8

ДИСЕРТАЦІЯ

**ІНТЕЛЕКТУАЛЬНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ В СИСТЕМІ
УПРАВЛІННЯ ВНЕСЕННЯМ ДОБРИВ**

122 «Комп'ютерні науки»
12 «Інформаційні технології»

Подається на здобуття ступеня доктора філософії

Дисертація містить результати власних досліджень.
Використання ідей, результатів і текстів інших авторів мають посилання
на відповідне джерело О. І. Качмарський

Науковий керівник
ГОЛУБ Белла Львівна,
кандидат технічних наук, доцент

Київ – 2026

АНОТАЦІЯ

Качмарський О. І. Інтелектуальні інформаційні технології в системі управління внесенням добрив. Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії за спеціальністю 122 «Комп'ютерні науки». Національний університет біоресурсів і природокористування України, Київ, 2026.

Дисертацію присвячено розробленню інтелектуальної технології підтримки прийняття рішень в умовах невизначеності, з прикладним застосуванням до прогнозування врожайності та оптимізації внесення добрив за прибутковістю з урахуванням ризику. Актуальність дослідження зумовлена тим, що клас задач прийняття рішень за умов одночасного впливу нелінійних залежностей між керованими параметрами та цільовою функцією, а також стохастичного впливу чинників середовища, потребує методологічного апарату, який поєднує прогнозні моделі машинного навчання з кількісною оцінкою ризику. Наявні інформаційні технології в цій предметній області здебільшого реалізують монолітну архітектуру, що не передбачає розмежування етапів навчання моделі та її експлуатації, що ускладнює роботу системи в режимі, наближеному до реального часу. Крім того, типові рекомендаційні алгоритми формують точкові оцінки оптимального рішення без аналізу розподілу можливих наслідків, що обмежує інформаційну цінність рекомендацій для особи, яка приймає рішення в умовах ризику. Розвиток методів машинного навчання та параметричних методів стохастичного моделювання відкриває можливості для подолання цих обмежень. Таким чином, актуальність дослідження полягає в розробленні архітектурного та математичного апарату, що інтегрує нелінійне прогнозування з ризик-орієнтованим аналізом у масштабовану програмну платформу, апробовану на прикладі оптимізації внесення мінеральних добрив.

Метою дисертації є розроблення моделей, алгоритмів та створення на їхній основі гібридної інтелектуальної інформаційної технології підтримки прийняття рішень, інтегрованої в загальну систему управління внесенням добрив для підвищення ефективності вирощування зернових культур.

Наукова новизна одержаних результатів полягає у наступному.

1. Розроблено алгоритм побудови інтегрованого інформаційного простору на основі агрономічної статистики та глобальних архівів метеорологічних даних, який відрізняється від чинних підходів тим, що включає широтно-залежний індикатор теплового стресу та процедуру формування проксі-показника продуктивності агровиробничої системи без витоку цільової змінної, і призначений для підвищення точності прогнозування врожайності в задачах підтримки прийняття рішень з нормування добрив.

2. Удосконалено модель підтримки прийняття рішень, яка відрізняється від чинних рішень тим, що інтегрує нелінійну прогнозну функцію відгуку врожайності з ризик-орієнтованим економічним функціоналом на сценарному просторі погодних реалізацій, і призначена для формування стійких до невизначеності стратегій внесення добрив замість точкових рекомендацій.

3. Вперше розроблено архітектуру системи підтримки прийняття рішень для управління внесенням добрив на основі інтеграції методів машинного навчання, математичного моделювання прибутковості з урахуванням ризиків та сучасних програмно-інженерних рішень: розроблені архітектурні зміни розширюють можливості вибору оптимальних доз добрив, забезпечують конвеєризацію обробки даних та кросплатформність ядра системи, дозволяючи розгортати її в різномірних клієнтських середовищах без модифікації прогнозної моделі.

Практичне значення отриманих результатів полягає у тому, що розроблена система може бути впроваджена у діяльність агровиробничих підприємств і фермерських господарств, агроконсультаційних служб для автоматизованої

підтримки прийняття рішень щодо норм внесення мінеральних добрив. Розроблений програмний прототип забезпечує формування адаптивних економічно обґрунтованих рекомендацій з кількісною оцінкою кліматичних і цінових ризиків, що сприяє підвищенню фінансової стійкості агровиробництва в умовах підвищеної мінливості погодних і ринкових чинників. Запропоновані методологічні та програмні рішення також можуть бути впроваджені у навчальний процес профільних спеціальностей закладів вищої освіти аграрного й комп'ютерного спрямування при викладанні дисциплін, що стосуються інтелектуального аналізу даних, систем підтримки прийняття рішень та інформаційних технологій в агропромисловому комплексі. Подальший розвиток прототипу до повноцінного виробничого рішення передбачає локалізацію набору даних до вузького географічного регіону зі схожими агрокліматичними умовами та дезагрегацію цільової змінної до рівня окремої зернової культури, інтеграцію ґрунтової складової простору ознак, а також підключення до зовнішніх джерел актуальних ринкових цін на добрива та сільськогосподарську продукцію через відповідні API.

Результати дослідження апробовано на міжнародних і всеукраїнських науково-практичних конференціях у галузі інформаційних технологій та агроінформатики. За тематикою дисертації опубліковано наукові праці у фахових виданнях України.

У вступі подано загальну характеристику дослідження, обґрунтовано актуальність теми, сформульовано мету та завдання, визначено об'єкт, предмет і методи дослідження, викладено наукову новизну та практичне значення отриманих результатів, наведено відомості про особистий внесок здобувача, апробацію результатів та структуру дисертації.

Перший розділ присвячено теоретико-методологічному аналізу проблеми створення інтелектуальних інформаційних технологій в управлінні внесенням добрив. Розглянуто еволюцію концепції точного землеробства, проаналізовано сучасні системи підтримки прийняття рішень в агрономії, методи

диференційованого внесення та комерційні аналітичні платформи. Проведено систематизацію інтелектуальних технологій в агросекторі за напрямками застосування: системи моніторингу на базі IoT, прогнозні моделі врожайності із застосуванням методів машинного навчання та алгоритми оптимізації агротехнологічних рішень. Здійснено класифікацію та порівняльний аналіз сучасних підходів до формування рекомендацій з нормування мінеральних добрив. Виявлено системну суперечність у сфері інформаційно-технологічного забезпечення точного землеробства: значний прогрес у засобах збору даних супроводжується відставанням алгоритмічного апарату прийняття рішень, який залишається переважно детермінованим та ігнорує стохастичну природу погодних ризиків і нелінійність біологічних процесів. Обґрунтовано методологічний розрив між наявністю точних прогнозних моделей та здатністю трансформувати ці прогнози в економічно стійкі стратегії управління. На основі проведеного аналізу сформульовано наукову задачу дослідження та визначено напрями її вирішення.

Другий розділ присвячено формалізації математичної моделі гібридної системи оптимізації доз добрив. Сформульовано задачу прийняття рішень як стохастичну оптимізаційну задачу, де цільова функція прибутку залежить від керованих змінних: вектора доз $u = (N, P, K)$ у просторі діючих речовин та випадкового вектора станів середовища, представленого агрегованими погодними індикаторами. Обґрунтовано вибір градієнтного бустингу регресійних дерев як прогнозного ядра, здатного коректно відтворювати ефекти насичення та перехресні взаємодії факторів. Математично формалізовано процедуру корекції зміщення прогнозу при переході від логарифмічного простору до натуральних величин з використанням властивостей логнормального розподілу. Розроблено алгоритм оптимізації за прибутковістю з урахуванням ризику на основі поєднання критерію Mean-Variance та емпіричного розрахунку CVaR на сценарному просторі. Виконано верифікацію моделі: на тестовій вибірці отримано $R^2 = 0.9122$, $RMSE = 581.48$ кг/га, $MAE = 376.77$ кг/га; п'ятикратна крос-

валідація показала $R^2 = 0.908 \pm 0.010$ та $RMSE = 599 \pm 24$ кг/га, що свідчить про стабільність узагальнювальної здатності.

Третій розділ присвячено архітектурній побудові та програмній реалізації прототипу інтелектуальної інформаційної технології. Запропоновано архітектуру з розділеними середовищами, що відокремлює дослідницький модуль машинного навчання (Python) від клієнтського настільного застосунку (.NET), які пов'язані серіалізованою у платформонезалежний формат прогнозною моделлю з її подальшим нативним інференсом у керованому коді клієнтського застосунку. Розроблено модульну структуру програмного комплексу з виділенням бібліотеки ядра (оркестратор CoreService, сервіси WeatherService, FertilizerService, PredictionService та ризик-орієнтований економічний блок), консольної утиліти пакетного збору даних та клієнтського застосунку з картографічним інтерфейсом для вибору географічної точки. Реалізовано механізми відмовостійкої взаємодії з відкритими API: повторні спроби з фіксованою паузою 60 с при HTTP 429 та лінійним збільшенням затримки при інших помилках, проміжне резервне копіювання при пакетному зборі даних.

Четвертий розділ містить результати обчислювальної апробації та оцінку ефективності розробленої технології. Виконано порівняльний аналіз запропонованої моделі з низкою альтернативних підходів до управління внесенням добрив: традиційною рекомендаційною схемою RDF, сенсорним точним азотним менеджментом, управлінням на основі машинного навчання з оптимізаційним пошуком та підходом за участі фермерів FPNM. Принциповим методологічним рішенням є оцінювання кожної рекомендації на тій географічній локації, де її було отримано: у Паралахемунді (Одіша, Індія), Педдапурамі (Андхра-Прадеш, Індія), Венсу (Сіньцзян, Китай) та Хенані (Китай), що забезпечує коректність порівняння інструментів підтримки рішень у тих самих агрокліматичних умовах. У межах кожної локації згенеровано $N = 500$ параметричних Монте-Карло-реалізацій погодних умов на основі 20-річної історії спостережень з Open-Meteo Archive API. Економічні розрахунки виконано

у трьох цінових сценаріях: низькому, помірному та кризовому. Запропонована система формує не одну точкову рекомендацію, а повну криву ризик-економічних показників у діапазоні доз 50-500 кг/га для чотирьох добрив (карбамід, аміачна селітра, NPK 15-15-15, NPK 20-20-20) з виділеними точками *Opt.MV* і *Opt.CVaR*. На локації Хенань зафіксовано стратегію з $CVaR_{10} = 10\,318$ грн/га, що у 3,46 рази перевищує $CVaR_{10}$ фіксованої рекомендації *Rep* (2 979 грн/га) при поступці в очікуваному прибутку 21,2%; на локації Педдапурам виявлено стратегії з переходом $CVaR_{10}$ з негативної в позитивну зону, недоступного з рекомендації *Sujatha*. Проведено абляційний аналіз структури вхідного простору ознак, який підтвердив, що контекстна ознака *CountryTechLevel* забезпечує приріст коефіцієнта детермінації з 0,807 до 0,9122 (+0,104), тоді як альтернативний трендовий регресор (календарний рік) дає значно менший приріст (+0,020) і свідомо відхилений як методологічно некоректний для задач *ex ante* прогнозування. Підтверджено інформаційну цінність системи як інструмента, що замінює точкову рекомендацію структурованим простором ризик-економічних компромісів з кількісно охарактеризованими наслідками кожної опції.

У висновках узагальнено результати дослідження та сформульовано основні наукові й практичні здобутки. Розроблено інтелектуальну інформаційну технологію, що поєднує методи машинного навчання, оптимізації за прибутковістю з урахуванням ризику та сучасні програмно-інженерні рішення для автоматизованої підтримки прийняття агрономічних рішень. Створено двокомпонентну програмну архітектуру, що інтегрує дослідницьке середовище Python/XGBoost з клієнтським .NET-застосунком через нативну інтеграцію серіалізованої моделі. За результатами обчислювальної апробації встановлено здатність запропонованої технології формувати стратегії з кращим ризиковим профілем порівняно з фіксованими рекомендаціями альтернативних підходів за рахунок кількісної характеристики розподілу економічних наслідків.

Ключові слова: інтелектуальна інформаційна технологія, система підтримки прийняття рішень, машинне навчання, штучний інтелект, гібридна система,

прийняття рішень в умовах невизначеності, прогнозування врожайності, градієнтний бустинг, оптимізація за прибутковістю з урахуванням ризику, стохастична оптимізація, математичне моделювання, інтелектуальне керування, точне землеробство, аграрний сектор, програмне забезпечення, автоматизація.

ANNOTATION

Kachmarskyi O. I. Intelligent information technologies in the system of fertilizer application management. The qualification scientific work on the rights of the manuscript.

The dissertation on acquisition of Doctor of Philosophy scientific degree in specialty 122 «Computer Science». National University of Life and Environmental Sciences of Ukraine, Kyiv, 2026.

The dissertation is devoted to the development of an intelligent technology for decision support under uncertainty, with an applied focus on yield forecasting and the optimization of fertilizer application with respect to profitability under risk consideration.

The relevance of the study is driven by the fact that the class of decision-making problems involving the simultaneous influence of nonlinear dependencies between controlled parameters and the objective function, together with the stochastic impact of environmental factors, requires a methodological apparatus that combines machine learning predictive models with quantitative risk assessment. Existing information technologies in this subject area predominantly implement a monolithic architecture that does not separate the model training stage from its operational deployment, which complicates system operation in a near real-time mode. Moreover, typical recommendation algorithms produce point estimates of the optimal decision without analyzing the distribution of possible outcomes, which limits the informational value of the recommendation for a decision-maker operating under risk. The development of machine learning methods and parametric stochastic modeling techniques opens new possibilities for overcoming these limitations. Thus, the relevance of the study lies in

the development of an architectural and mathematical apparatus that integrates nonlinear forecasting with risk-oriented analysis into a scalable software platform, validated on the example of mineral fertilizer application optimization.

The aim of the dissertation is to develop models, algorithms, and, on their basis, a hybrid intelligent information technology for decision support, integrated into the overall fertilizer application management system, in order to improve the efficiency of grain crop cultivation.

The scientific novelty of the obtained results is as follows.

1. An algorithm for constructing an integrated information space based on agronomic statistics and global meteorological archives has been developed, which differs from existing approaches in that it includes a latitude-dependent indicator of heat stress and a procedure for forming a proxy indicator of agricultural production system productivity without leakage of the target variable, and is intended to improve the accuracy of yield forecasting in decision support tasks for fertilizer rate determination.
2. A decision support model has been improved, which differs from existing solutions in that it integrates a nonlinear predictive yield response function with a risk-oriented economic functional over the scenario space of weather realizations, and is intended for forming strategies for fertilizer application that are robust to uncertainty, as opposed to point recommendations.
3. For the first time, an architecture for a decision support system for fertilizer application management has been developed, based on the integration of machine learning methods, mathematical modeling of profitability under risk, and modern software engineering solutions: the proposed architectural changes expand the possibilities for selecting optimal fertilizer doses, ensure the pipelining of data processing, and provide cross-platform capability of the system core, allowing its deployment across heterogeneous client environments without modifying the predictive model.

The practical significance of the obtained results lies in the fact that the developed system can be implemented in the activities of agricultural enterprises and farm holdings, as well as agricultural advisory services, for the automated decision support regarding mineral fertilizer application rates. The developed software prototype enables the formation of adaptive, economically justified recommendations with quantitative assessment of climatic and price risks, thereby contributing to the improved financial resilience of agricultural production under conditions of increased volatility of weather and market factors. The proposed methodological and software solutions can also be implemented in the educational process of relevant agricultural and computer science specialties at higher education institutions, in teaching disciplines related to intelligent data analysis, decision support systems, and information technologies in the agro-industrial complex. Further development of the prototype into a fully-fledged production solution involves localizing the dataset to a narrower geographic region with similar agroclimatic conditions, disaggregating the target variable to the level of an individual grain crop, integrating a soil-related component into the feature space, as well as connecting to external sources of up-to-date market prices for fertilizers and agricultural products via corresponding APIs.

The results of the study have been presented at international and national scientific-practical conferences in the fields of information technology and agroinformatics. Scientific papers on the topic of the dissertation have been published in professional journals of Ukraine.

The introduction provides a general characterization of the study, substantiates the relevance of the topic, formulates the aim and objectives, defines the object, subject, and methods of the research, presents the scientific novelty and practical significance of the obtained results, and provides information on the personal contribution of the candidate, the validation of the results, and the structure of the dissertation.

The first chapter is devoted to a theoretical and methodological analysis of the problem of creating intelligent information technologies for fertilizer application management. The evolution of the precision agriculture concept is examined, modern

decision support systems in agronomy, variable-rate application methods, and commercial analytical platforms are analyzed. A systematization of intelligent technologies in the agricultural sector has been carried out by application areas: IoT-based monitoring systems, yield prediction models using machine learning methods, and algorithms for optimizing agrotechnological decisions. A classification and comparative analysis of modern approaches to forming mineral fertilizer application rate recommendations has been performed. A systemic contradiction in the field of information-technological support for precision agriculture has been revealed: significant progress in data collection tools is accompanied by a lag in the algorithmic apparatus of decision-making, which remains predominantly deterministic and disregards the stochastic nature of weather risks and the nonlinearity of biological processes. A methodological gap between the availability of accurate predictive models and the ability to transform these forecasts into economically resilient management strategies has been substantiated. Based on the conducted analysis, the scientific task of the study has been formulated and the directions for its resolution have been identified.

The second chapter is devoted to the formalization of the mathematical model of the hybrid fertilizer dose optimization system. The decision-making problem is formulated as a stochastic optimization problem, where the profit objective function depends on controlled variables — the dose vector $u = (N, P, K)$ in the space of active substances — and a random vector of environmental states, represented by aggregated weather indicators. The choice of gradient boosting of regression trees as the predictive core, capable of correctly reproducing saturation effects and cross-factor interactions, is substantiated. The procedure for bias correction of the forecast when transitioning from the logarithmic space to natural values, using the properties of the lognormal distribution, has been mathematically formalized. An algorithm for profitability optimization under risk consideration has been developed, based on the combination of the Mean-Variance criterion and the empirical calculation of CVaR over the scenario space. Model verification was performed: on the test set, $R^2 = 0.9122$, $RMSE = 581.48$ kg/ha, $MAE = 376.77$ kg/ha were obtained; five-fold cross-validation showed $R^2 =$

0.908 ± 0.010 and $RMSE = 599 \pm 24$ kg/ha, indicating the stability of the generalization capability.

The third chapter is devoted to the architectural design and software implementation of the intelligent information technology prototype. An architecture with separated environments is proposed, separating the machine learning research module (Python) from the client desktop application (.NET), which are connected by a predictive model serialized into a platform-independent format, with its subsequent native inference in the managed code of the client application. A modular structure of the software complex has been developed, with the core library (the CoreService orchestrator, the WeatherService, FertilizerService, PredictionService services, and the risk-oriented economic module) singled out, along with a console utility for batch data collection and a client application with a cartographic interface for selecting a geographic point. Fault-tolerant interaction mechanisms with open APIs have been implemented: retry attempts with a fixed 60-second pause on HTTP 429 responses and linearly increasing delay on other errors, as well as intermediate backup during batch data collection.

The fourth chapter contains the results of the computational validation and efficiency evaluation of the developed technology. A comparative analysis of the proposed model against a range of alternative approaches to fertilizer application management was performed: the traditional recommendation scheme RDF, sensor-based precision nitrogen management, machine-learning-based management with optimization search, and the farmer-participation approach FPNM. A fundamental methodological decision was to evaluate each recommendation at the geographic location where it was originally obtained: in Paralakhemundi (Odisha, India), Peddapuram (Andhra Pradesh, India), Wensu (Xinjiang, China), and Henan (China), which ensures the correctness of comparing decision support tools under the same agroclimatic conditions. Within each location, $N = 500$ parametric Monte Carlo realizations of weather conditions were generated based on a 20-year history of observations from the Open-Meteo Archive API. Economic calculations were

performed under three price scenarios: low, moderate, and crisis. The proposed system forms not a single point recommendation, but a full curve of risk-economic indicators across a dose range of 50–500 kg/ha for four fertilizers (urea, ammonium nitrate, NPK 15-15-15, NPK 20-20-20), with highlighted Opt.MV and Opt.CVaR points. At the Henan location, a strategy with $CVaR_{10} = 10,318$ UAH/ha was identified, which is 3.46 times higher than the $CVaR_{10}$ of the fixed Ren recommendation (2,979 UAH/ha), at a concession in expected profit of 21.2%; at the Peddapuram location, strategies were identified with a transition of $CVaR_{10}$ from negative to positive territory, unattainable with the Sujatha recommendation. An ablation analysis of the input feature space structure was conducted, confirming that the contextual feature CountryTechLevel provides an increase in the coefficient of determination from 0.807 to 0.9122 (+0.104), whereas the alternative trend regressor (calendar year) yields a substantially smaller increase (+0.020) and was deliberately rejected as methodologically inappropriate for ex ante forecasting tasks. The informational value of the system as a tool replacing a point recommendation with a structured space of risk-economic trade-offs, with quantitatively characterized consequences for each option, has been confirmed.

The conclusions summarize the results of the study and formulate the main scientific and practical achievements. An intelligent information technology has been developed that combines machine learning methods, profitability optimization under risk consideration, and modern software engineering solutions for the automated support of agronomic decision-making. A two-component software architecture has been created, integrating a Python/XGBoost research environment with a .NET client application through the native integration of a serialized model. Based on the results of the computational validation, the ability of the proposed technology to form strategies with a more favorable risk profile compared to the fixed recommendations of alternative approaches has been established, owing to the quantitative characterization of the distribution of economic outcomes under stochastic weather and price scenarios.

Keywords: intelligent information technology, decision support system, machine learning, artificial intelligence, hybrid system, decision-making under

uncertainty, yield forecasting, gradient boosting, profitability optimization under risk consideration, stochastic optimization, mathematical modeling, intelligent control, precision agriculture, agricultural sector, software, automation.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Статті у наукових виданнях, включених до Переліку наукових фахових видань України

1. Качмарський О. І., Голуб Б. Л. Аналіз стану впровадження сучасних інформаційних технологій у сільському господарстві. Наука і техніка сьогодні (категорія “Б”). 2025. № 8(49). С. 1465-1478.

2. Качмарський О. І., Голуб Б. Л., Панчук Т. В. Гібридна система підтримки рішень для економічної оптимізації доз добрив за змінних погодних умов. Наука і техніка сьогодні (категорія “Б”). 2025. № 11(52). С. 2236-2244.

3. Качмарський О. І., Голуб Б. Л. Математичне моделювання економіко-ризикової оптимізації доз мінеральних добрив у стохастичних погодних умовах з використанням моделей машинного навчання. Таврійський науковий вісник. Серія: Технічні науки, 1(1), с. 113–121.

4. Качмарський О. І., Голуб Б. Л., Програмна реалізація та експериментальна апробація гібридної системи економіко-ризикової оптимізації доз мінеральних добрив. Наука і техніка сьогодні (категорія “Б”). 2026. № 4(58). С. 3396-3416.

Перелік тез наукових доповідей.

5. Качмарський О. І. Використання машинного зору в сучасній агропромисловості. XIV Міжнародна науково-практична конференція молодих вчених «Інформаційні технології: економіка, техніка, освіта ‘2023», м Київ, 26-27 жовтня 2023 року. С. 77-78.

6. Качмарський О. І. Визначення стадій зрілості полуниці за допомогою згорткової нейронної мережі. XII Міжнародна науково-практична конференція

“Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні 2024”, м. Київ, 21-22 листопада 2024 року.

7. Качмарський О. І. Визначення фази росту пшениці за шкалою ВВСН за допомогою комп’ютерного зору. Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення, м. Тернопіль, Україна, м. Ополе, Польща, 13-14 березня 2025 року. С. 17-19.

8. Качмарський О. І., Голуб Б. Л. Метод класифікації фаз розвитку пшениці за допомогою комп’ютерного зору. VII науково-практична конференція студентів і аспірантів «Теоретичні та прикладні аспекти розробки комп’ютерних систем 2025», м. Київ, 24 квітня 2025 року. С. 136-137.

9. Качмарський О. І., Голуб Б. Л. Гібридна система підтримки рішень для економічної оптимізації доз добрив у змінних погодних умовах. XVI Міжнародна науково-практична конференція молодих вчених “Інформаційні технології: економіка, техніка, освіта”, м. Київ, 28-29 жовтня 2025 року. С. 308-309.

10. Качмарський О. І., Голуб Б. Л. Гібридна система підтримки аграрних рішень з економіко-ризиковою оптимізацією доз мінеральних добрив. VII всеукраїнська науково-практична конференція студентів і аспірантів “Теоретичні та прикладні аспекти розробки комп’ютерних систем 2026”, м. Київ, 29 квітня 2026 року.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	18
ВСТУП	20
РОЗДІЛ 1 ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ЗАСТОСУВАННЯ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ У ПРОЦЕСАХ ВНЕСЕННЯ ДОБРИВ ...	31
1.1 Аналіз сучасних систем точного землеробства та методів управління живленням культур	31
1.2 Інтелектуальні інформаційні технології в агрономії (підходи, архітектури, аналітичні моделі)	39
1.3. Проблеми адаптації моделей прийняття рішень до змінних природно-кліматичних факторів	44
1.4. Узагальнення теоретичних положень та постановка наукового завдання ..	55
Висновки до розділу 1	55
РОЗДІЛ 2 МАТЕМАТИЧНА МОДЕЛЬ ГІБРИДНОЇ СИСТЕМИ ОПТИМІЗАЦІЇ ДОЗ ДОБРИВ	58
2.1. Формалізація задачі оптимізації доз мінеральних добрив у стохастичних умовах	58
2.2. Побудова прогнозової моделі врожайності на основі алгоритмів машинного навчання	68
2.3. Верифікація та метричний аналіз точності запропонованої моделі	73
Висновки до розділу 2	85
РОЗДІЛ 3 РЕАЛІЗАЦІЯ СИСТЕМИ ПІДТРИМКИ РІШЕНЬ ДЛЯ УПРАВЛІННЯ ВНЕСЕННЯМ ДОБРИВ З ІНТЕЛЕКТУАЛЬНОЮ СКЛАДОВОЮ	87
3.1. Архітектурна структура та взаємодія модулів гібридної системи	87
3.2. Алгоритмічне забезпечення агрономічного, прогнозного та економічного компонентів	97
3.3. Реалізація прототипу інтелектуальної системи на базі інтегрованих даних про погоду	107
3.4. Інтерпретація результатів та оцінка ефективності	124
Висновки до розділу 3	129
РОЗДІЛ 4 ОБЧИСЛЮВАЛЬНА АПРОБАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ РОЗРОБЛЕНОЇ ТЕХНОЛОГІЇ	131

4.1. Верифікація прогнозової моделі на даних інших досліджень	131
4.2. Економічна ефективність та ризик-орієнтована оцінка рішень у різних погодних сценаріях	134
4.3. Узагальнення результатів обчислювальної апробації та практичні рекомендації щодо застосування системи	148
Висновки до розділу 4	163
ВИСНОВКИ	165
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	169
ДОДАТКИ	180

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

СППР – система підтримки прийняття рішень.

ML – Machine Learning, машинне навчання.

API – Application Programming Interface, програмний інтерфейс взаємодії.

WPF – Windows Presentation Foundation, технологія побудови графічних інтерфейсів на платформі .NET.

НРК – комплекс макроелементів живлення рослин (азот, фосфор, калій).

FAOSTAT – статистична база даних Продовольчої і сільськогосподарської організації ООН (FAO), що містить агростатистичні показники за країнами і роками.

Open-Meteo – відкритий сервіс історичних та поточних метеоданих.

XGBoost – алгоритм градієнтного бустингу регресійних дерев.

RF – Random Forest, алгоритм випадкового лісу.

LR – Linear Regression, лінійна регресія.

CV – Cross-Validation, перехресна валідація (за замовчуванням п'ятикратна).

R² – коефіцієнт детермінації.

MSE, RMSE – Mean Squared Error і Root Mean Squared Error, середньоквадратична похибка та її корінь, кг/га.

MAE, MedAE – Mean Absolute Error і Median Absolute Error, середня та медіанна абсолютна похибка, кг/га.

MAPE, sMAPE, WAPE – Mean Absolute Percentage Error, symmetric MAPE та Weighted Absolute Percentage Error, варіації відсоткової похибки прогнозу, %.

Mean-Variance – підхід до скаляризації багатокритеріальної оптимізації, що поєднує математичне сподівання та дисперсію цільової функції.

VaR α – Value-at-Risk на рівні α , α -квантиль розподілу прибутку.

CVaR α – Conditional Value-at-Risk, математичне сподівання прибутку за умови, що він гірший за VaR α .

CV (коеф. варіації) – коефіцієнт варіації прибутку, %.

Leak-free – режим обчислення ознак, що виключає інформаційний витік

цільової змінної з тренувальної підвибірki у тестову.

Абляційний аналіз – метод обґрунтування структури вхідного простору ознак шляхом поетапного додавання компонентів і вимірювання приросту якості моделі.

ВСТУП

Обґрунтування вибору теми дослідження. Сільськогосподарське виробництво функціонує в умовах високої невизначеності, зумовленої стохастичною природою погодних факторів, біологічною варіабельністю рослин та динамікою ринкових цін на продукцію та засоби виробництва. В умовах зростання населення планети, обмеженості земельних ресурсів та зростаючої кліматичної мінливості традиційні екстенсивні методи господарювання вичерпують свій потенціал. В Україні це питання набуває ще більшої гостроти у зв'язку з військовими діями. Підвищення ефективності агровиробництва, зокрема управління мінеральним живленням рослин, вимагає переходу до концепції точного землеробства, де ключовим активом стають дані, а інструментом конкурентної боротьби – інтелектуальні алгоритми їх обробки.

Внесення мінеральних добрив є однією з найбільш капіталомістких операцій у структурі прямих витрат рослинництва, складаючи 30–50% витрат на гектар. Існуючі підходи до розрахунку доз добрив, що базуються на нормативних методах і балансових рівняннях, часто оперують усередненими багаторічними показниками і не враховують ані специфіки конкретного року, ані поточної цінової кон'юнктури. Як наслідок виникають приховані фінансові втрати: у сприятливі сезони не використовується продуктивний потенціал агроценозу, у несприятливі фермер несе підвищені витрати на добрива при низькій реалізованій врожайності.

Розвиток методів штучного інтелекту, зокрема, ансамблевих алгоритмів машинного навчання та методів стохастичної оптимізації, відкриває нові можливості для виявлення складних нелінійних залежностей між агротехнологічними факторами та продуктивністю посівів. Водночас наявність точного прогнозу врожайності сама по собі не гарантує ефективності управлінського рішення. Виникає потреба у створенні комплексних інтелектуальних інформаційних технологій, що здатні не лише прогнозувати стан агроценозу, а й трансформувати прогнози в економічно обґрунтовані рекомендації з кількісним урахуванням ризиків, та забезпечують безшовну

інтеграцію наукових методів у виробниче програмне середовище агропідприємств. Таким чином, актуальність розроблення інтелектуальних інформаційних технологій в системі управління внесенням добрив зумовлена як глобальними викликами цифрової трансформації агросектору, так і потребою в подоланні методологічного розриву між точністю прогнозних моделей та якістю управлінських рішень в умовах кліматичної та цінової невизначеності.

Зв'язок роботи з науковими програмами, планами, темами, грантами.

Дисертаційне дослідження виконано в руслі пріоритетних напрямів розвитку науки та техніки України, що стосуються впровадження інтелектуальних інформаційних систем в агрономічній галузі. Робота концептуально узгоджується з цілями наукових програм та стратегічних ініціатив у галузі сільського господарства, сталого розвитку та цифрової трансформації державного управління.

Дослідження виконані відповідно до планів науково-дослідницької роботи Національного університету біоресурсів і природокористування за темою «Інформаційна та алгоритмічна підтримка інтелектуальних систем в природоохоронній галузі» (номер державної реєстрації №0124U002754).

Мета дослідження. Метою дисертаційної роботи є розроблення моделей, алгоритмів та створення на їхній основі гібридної інтелектуальної інформаційної технології підтримки прийняття рішень, інтегрованої в загальну систему управління внесенням добрив для підвищення ефективності вирощування зернових культур.

Завдання дослідження. Для досягнення поставленої мети сформульовано такі завдання:

1. Провести системний аналіз сучасних методів управління внесенням мінеральних добрив, інтелектуальних інформаційних технологій в агрономії та підходів до прогнозування врожайності й виявити обмеження традиційних детермінованих підходів в умовах стохастичної невизначеності погодних і цінових факторів.

2. Сформувати інтегрований інформаційний простір на основі відкритих джерел зі структурою ознак, що охоплює агрономічні та сезонні кліматичні індикатори, а також контекстний регресор CountryTechLevel як проксі-показник сукупної продуктивності агровиробничої системи країни.

3. Побудувати прогнозну модель врожайності на основі ансамблевого методу градієнтного бустингу в логарифмічному просторі цільової змінної з аналітичною корекцією логнормального зміщення, виконати її верифікацію на тестовій вибірці та формалізований абляційний аналіз структури вхідного простору ознак.

4. Розробити математичну модель оптимізації внесення добрив за прибутковістю з урахуванням ризику, що інтегрує нелінійну прогнозну функцію відгуку з ризик-орієнтованим економічним функціоналом на основі поєднання Mean-Variance скаляризації та емпіричного розрахунку Conditional Value-at-Risk (CVaR) на сценарному просторі погодних і цінових реалізацій.

5. Спроекувати та програмно реалізувати двокомпонентну архітектуру інтелектуальної технології, що відокремлює дослідницьке середовище машинного навчання (Python) від клієнтського застосунку бізнес-логіки (.NET) через серіалізацію прогновної моделі у платформонезалежний формат з її нативним інференсом у керованому коді .NET-платформи, а також містить модуль сценарної оптимізації за прибутковістю з урахуванням ризику.

6. Виконати обчислювальну апробацію розробленої технології на сценарному просторі погодних реалізацій, провести порівняльний аналіз з традиційними та сучасними альтернативними підходами та оцінити стійкість оптимізованих рекомендацій до макроекономічних і кліматичних збурень.

Об'єкт дослідження. Об'єктом дослідження є процеси інтелектуального аналізу агрономічних і кліматичних даних та комп'ютерного моделювання в системах підтримки прийняття рішень для управління внесенням добрив при вирощуванні зернових культур.

Предмет дослідження. Предметом дослідження є застосування методів машинного навчання, математичного моделювання та фінансової оптимізації в системах підтримки прийняття рішень.

Методи дослідження. У роботі застосовано комплекс методів, що забезпечують системний підхід до проєктування та реалізації інтелектуальних інформаційних технологій автоматизованої підтримки прийняття рішень в агрономії.

У дисертації застосовано методи:

- 1) системного аналізу для формалізації задачі та проєктування дворівневої архітектури системи;
- 2) інтеграції з відкритими джерелами даних для збору агрономічної статистики та метеорологічних рядів з механізмами відмовостійкої мережевої взаємодії;
- 3) машинного навчання для побудови нелінійної прогнозовної моделі врожайності методом ансамблевого градієнтного бустингу в логарифмічному просторі цільової змінної;
- 4) обґрунтування структури простору ознак методом абляційного аналізу з п'ятикратною перехресною перевіркою на відкладеній тестовій вибірці;
- 5) багатокритеріальної оптимізації з урахуванням ризику на основі Mean-Variance скаляризації та емпіричного розрахунку Conditional Value-at-Risk;
- 6) об'єктно-орієнтованого проєктування для побудови модульної архітектури програмного комплексу;
- 7) математичної статистики для оцінювання значущості обчислювальної апробації результатів.

Наукова новизна одержаних результатів полягає у наступному:

- Вперше розроблено алгоритм побудови інтегрованого інформаційного простору на базі агрономічної статистики та глобальних архівів метеорологічних даних, що охоплює комплекс агрономічних, кліматичних

індикаторів і проксі-показник потенціалу агровиробничої системи країни. Це забезпечило оптимальне формування ознакового простору в контурі управління внесенням добрив.

- Удосконалено модель оптимізації норм внесення мінеральних добрив шляхом інтеграції нелінійної прогнозної функції відгуку врожайності з оптимізацією за прибутковістю з урахуванням ризику, що дозволило сформувати стійкі до невизначеності стратегії внесення добрив замість класичних точкових рекомендацій в умовах погодної та ринкової невизначеності.

- Вперше розроблено архітектуру системи підтримки прийняття рішень для управління внесенням добрив на основі інтеграції методів машинного навчання, математичного моделювання прибутковості з урахуванням ризиків та сучасних програмно-інженерних рішень. Розроблені архітектурні зміни розширюють можливості вибору оптимальних доз добрив, забезпечують конвеєризацію обробки даних та кросплатформність ядра системи, дозволяючи розгортати її в різномірних клієнтських середовищах без модифікації прогнозної моделі.

Практичне значення отриманих результатів. Практичне значення отриманих результатів полягає у тому, що розроблена інтелектуальна інформаційна технологія може бути впроваджена у діяльність агровиробничих підприємств і фермерських господарств, агроконсультаційних служб для автоматизованої підтримки прийняття рішень щодо норм внесення мінеральних добрив. Розроблений програмний прототип забезпечує формування адаптивних економічно обґрунтованих рекомендацій з кількісною оцінкою кліматичних і цінових ризиків, що сприяє підвищенню фінансової стійкості агровиробництва в умовах підвищеної мінливості погодних і ринкових чинників. Запропоновані методологічні та програмні рішення також можуть бути впроваджені у навчальний процес профільних спеціальностей закладів вищої освіти аграрного й комп'ютерного спрямування при викладанні дисциплін, що стосуються інтелектуального аналізу даних, систем підтримки прийняття рішень та

інформаційних технологій в агропромисловому комплексі. Подальший розвиток прототипу до повноцінного виробничого рішення передбачає локалізацію набору даних до вузького географічного регіону зі схожими агрокліматичними умовами та дезагрегацію цільової змінної до рівня окремої зернової культури, інтеграцію ґрунтової складової простору ознак, а також підключення до зовнішніх джерел актуальних ринкових цін на добрива та сільськогосподарську продукцію через відповідні API.

Особистий внесок здобувача. Усі основні наукові результати, що виносяться на захист, отримано здобувачем самостійно. У дисертаційній роботі автором особисто виконано: системний аналіз літературних джерел та формулювання наукової задачі дослідження; формування інтегрованого інформаційного простору на основі відкритих джерел FAOSTAT та Open-Meteo Archive API; розроблення структури вхідного простору ознак, включно з власними агрокліматичними індикаторами (GDD, SDD з широтно-залежним порогом) та контекстним регресором CountryTechLevel; побудову та верифікацію прогнозної моделі врожайності на основі XGBoost з аналітичною корекцією логнормального зміщення; формалізований абляційний аналіз структури вхідного простору ознак; розроблення алгоритму оптимізації за прибутковістю з урахуванням ризику на основі поєднання Mean-Variance та CVaR; проєктування та програмну реалізацію модульної архітектури програмного комплексу (бібліотека ядра, консольна утиліта збору даних, клієнтський застосунок); інтеграцію прогнозної моделі у програмний код клієнтського застосунку; реалізацію механізмів відмовостійкої взаємодії з відкритими API; проведення обчислювальної апробації та інтерпретацію результатів. У спільних публікаціях зі співавторами особистий внесок здобувача полягає у формуванні математичних постановок, розробленні алгоритмічних рішень, програмній реалізації, проведенні обчислювальної апробації, обробці та інтерпретації результатів, підготовці тексту рукописів. Науковий керівник здійснював концептуальне

керівництво, постановку наукових завдань, валідацію методологічних рішень та редакційне доопрацювання рукописів.

Апробація результатів дослідження. Основні результати дисертаційного дослідження доповідалися та обговорювалися на міжнародних і всеукраїнських науково-практичних конференціях у галузі інформаційних технологій.

- XIV Міжнародна науково-практична конференція молодих вчених «Інформаційні технології: економіка, техніка, освіта '2023», 26–27 жовтня 2023 року, НУБіП України, Київ.

- XII Міжнародна науково-практична конференція «Глобальні та регіональні проблеми інформатизації в суспільстві і природокористуванні 2024», 21–22 листопада 2024 року, Київ.

- Міжнародна наукова інтернет-конференція «Інформаційне суспільство: технологічні, економічні та технічні аспекти становлення», 13–14 березня 2025 року, Тернопіль (Україна) – Ополе (Польща).

- VII науково-практична конференція студентів і аспірантів «Теоретичні та прикладні аспекти розробки комп'ютерних систем 2025», 24 квітня 2025 року, НУБіП України, Київ.

- XVI Міжнародна науково-практична конференція молодих вчених «Інформаційні технології: економіка, техніка, освіта», 28–29 жовтня 2025 року, Київ.

- VII Всеукраїнська науково-практична конференція студентів і аспірантів «Теоретичні та прикладні аспекти розробки комп'ютерних систем 2026», 29 квітня 2026 року, НУБіП України, Київ.

Публікації. За тематикою дисертації опубліковано наукові праці у фахових виданнях.

- Качмарський О. І., Голуб Б. Л. Аналіз стану впровадження сучасних інформаційних технологій у сільському господарстві. Наука і техніка сьогодні (категорія “Б”). 2025. № 8(49). С. 1465-1478.

- Качмарський О. І., Голуб Б. Л., Панчук Т. В. Гібридна система підтримки рішень для економічної оптимізації доз добрив за змінних погодних умов. Наука і техніка сьогодні (категорія “Б”). 2025. № 11(52). С. 2236-2244.
- Качмарський О. І., Голуб Б. Л. Математичне моделювання економіко-ризикової оптимізації доз мінеральних добрив у стохастичних погодних умовах з використанням моделей машинного навчання. Таврійський науковий вісник. Серія: Технічні науки, 1(1), с. 113–121.
- Качмарський О. І., Голуб Б. Л., Програмна реалізація та експериментальна апробація гібридної системи економіко-ризикової оптимізації доз мінеральних добрив програмна реалізація та експериментальна апробація гібридної системи економіко-ризикової оптимізації доз мінеральних добрив . Наука і техніка сьогодні (категорія “Б”). 2026. № 4(58). С. 3396-3416.

Структура дисертації. Дисертаційна робота складається зі вступу, чотирьох розділів, висновків, списку використаних джерел та додатків. Загальний обсяг дисертації становить 201 сторінку. Робота містить 25 рисунків, 20 таблиць та 107 найменувань у списку використаних джерел.

Перший розділ присвячено теоретико-методологічному аналізу проблеми створення інтелектуальних інформаційних технологій в управлінні внесенням добрив. Розглянуто еволюцію концепції точного землеробства, проаналізовано сучасні системи підтримки прийняття рішень в агрономії, методи диференційованого внесення та комерційні аналітичні платформи. Проведено систематизацію інтелектуальних технологій в агросекторі за напрямками застосування: системи моніторингу на базі IoT, прогнознi моделі врожайності із застосуванням методів машинного навчання та алгоритми оптимізації агротехнологічних рішень. Здійснено класифікацію та порівняльний аналіз сучасних підходів до формування рекомендацій з нормування мінеральних добрив. Виявлено системну суперечність у сфері інформаційно-технологічного забезпечення точного землеробства: значний прогрес у засобах збору даних супроводжується відставанням алгоритмічного апарату прийняття рішень, який

залишається переважно детермінованим та ігнорує стохастичну природу погодних ризиків і нелінійність біологічних процесів. Обґрунтовано методологічний розрив між наявністю точних прогнозних моделей та здатністю трансформувати ці прогнози в економічно стійкі стратегії управління. На основі проведеного аналізу сформульовано наукову задачу дослідження та визначено напрями її вирішення.

Другий розділ присвячено формалізації математичної моделі гібридної системи оптимізації доз добрив. Сформульовано задачу прийняття рішень як стохастичну оптимізаційну задачу, де цільова функція прибутку залежить від керованих змінних – вектора доз $u = (N, P, K)$ у просторі діючих речовин – та випадкового вектора станів середовища, представленого агрегованими погодними індикаторами. Обґрунтовано вибір градієнтного бустингу регресійних дерев як прогнозного ядра, здатного коректно відтворювати ефекти насичення та перехресні взаємодії факторів. Математично формалізовано процедуру корекції зміщення прогнозу при переході від логарифмічного простору до натуральних величин з використанням властивостей логнормального розподілу. Розроблено алгоритм оптимізації за прибутковістю з урахуванням ризику на основі поєднання критерію Mean-Variance та емпіричного розрахунку CVaR на сценарному просторі. Виконано верифікацію моделі: на тестовій вибірці отримано $R^2 = 0.9122$, $RMSE = 581.48$ кг/га, $MAE = 376.77$ кг/га; п'ятикратна крос-валідація показала $R^2 = 0.908 \pm 0.010$ та $RMSE = 599 \pm 24$ кг/га, що свідчить про стабільність узагальнювальної здатності.

Третій розділ присвячено архітектурній побудові та програмній реалізації прототипу інтелектуальної інформаційної технології. Запропоновано архітектуру з розділеними середовищами, що відокремлює дослідницький модуль машинного навчання (Python) від клієнтського настільного застосунку (.NET), які пов'язані серіалізованою у платформонезалежний формат прогнозного моделлю з її подальшим нативним інференсом у керованому коді клієнтського застосунку. Розроблено модульну структуру програмного комплексу з виділенням бібліотеки

ядра (оркестратор CoreService, сервіси WeatherService, FertilizerService, PredictionService та ризик-орієнтований економічний блок), консольної утиліти пакетного збору даних та клієнтського застосунку з картографічним інтерфейсом для вибору географічної точки. Реалізовано механізми відмовостійкої взаємодії з відкритими API: повторні спроби з фіксованою паузою 60 с при HTTP 429 та лінійним збільшенням затримки при інших помилках, проміжне резервне копіювання при пакетному зборі даних.

Четвертий розділ містить результати обчислювальної апробації та оцінку ефективності розробленої технології. Виконано порівняльний аналіз запропонованої моделі з низкою альтернативних підходів до управління внесенням добрив: традиційною рекомендаційною схемою RDF, сенсорним точним азотним менеджментом, управлінням на основі машинного навчання з оптимізаційним пошуком та підходом за участі фермерів FPNM. Принциповим методологічним рішенням є оцінювання кожної рекомендації на тій географічній локації, де її було отримано — у Паралахемунді (Одіша, Індія), Педдапурамі (Андхра-Прадеш, Індія), Венсу (Сіньцзян, Китай) та Хенані (Китай), що забезпечує коректність порівняння інструментів підтримки рішень у тих самих агрокліматичних умовах. У межах кожної локації згенеровано $N = 500$ параметричних Монте-Карло-реалізацій погодних умов на основі 20-річної історії спостережень з Open-Meteo Archive API. Економічні розрахунки виконано у трьох цінових сценаріях — низькому, помірному та кризовому. Запропонована система формує не одну точкову рекомендацію, а повну криву ризик-економічних показників у діапазоні доз 50-500 кг/га для чотирьох добрив (карбамід, аміачна селітра, NPK 15-15-15, NPK 20-20-20) з виділеними точками *Opt.MV* і *Opt.CVaR*. На локації Хенань зафіксовано стратегію з $CVaR_{10} = 10\,318$ грн/га, що у 3,46 рази перевищує $CVaR_{10}$ фіксованої рекомендації Rep (2 979 грн/га) при поступці в очікуваному прибутку 21,2%; на локації Педдапурам виявлено стратегії з переходом $CVaR_{10}$ з негативної в позитивну зону, недоступного з рекомендації Sujatha. Проведено абляційний аналіз структури вхідного простору ознак, який

підтвердив, що контекстна ознака CountryTechLevel забезпечує приріст коефіцієнта детермінації з 0,807 до 0,9122 (+0,104), тоді як альтернативний трендовий регресор (календарний рік) дає значно менший приріст (+0,020) і свідомо відхилений як методологічно некоректний для задач *ex ante* прогнозування. Підтверджено інформаційну цінність системи як інструмента, що замінює точкову рекомендацію структурованим простором ризик-економічних компромісів з кількісно охарактеризованими наслідками кожної опції.

РОЗДІЛ 1

ТЕОРЕТИКО-МЕТОДОЛОГІЧНІ ОСНОВИ ЗАСТОСУВАННЯ ІНТЕЛЕКТУАЛЬНИХ ТЕХНОЛОГІЙ У ПРОЦЕСАХ ВНЕСЕННЯ ДОБРИВ

1.1 Аналіз сучасних систем точного землеробства та методів управління живленням культур

Сучасний агропромисловий комплекс функціонує в умовах обмеженості ресурсів та кліматичної нестабільності, що актуалізує проблему забезпечення продовольчої безпеки. Сільське господарство традиційно належить до найризикованіших галузей економіки. Виробничий процес відбувається у відкритій стохастичній системі та безпосередньо залежить від некерованих природних чинників, зокрема температурного режиму, розподілу опадів, динаміки розвитку патогенів і варіабельності ґрунтових характеристик. [22].

За прогнозами ООН, до 2050 року чисельність населення Землі досягне позначки у 9.7 мільярда осіб. Щоб задовольнити потреби такого населення, обсяг виробництва продовольства має зрости на 60–70%, тоді як можливості розширення площ орних земель практично вичерпано і, за оцінками, можуть зрости не більше, ніж на 4%. Крім цього, інтенсивна експлуатація земельних ресурсів призводить до деградації ґрунтів, зниження їх родючості та порушення екологічної рівноваги.

Цей дисбаланс між попитом на продовольство та наявними ресурсами до переходу від екстенсивних методів господарювання до парадигми – «Сільське господарство 4.0». В основі цієї концепції лежить інтеграція кіберфізичних систем, Інтернету речей (IoT), штучного інтелекту (ШІ) та аналітики великих даних (Big Data) у єдиний контур управління, що дозволяє перейти від реактивного реагування на проблеми до проактивного та предиктивного управління виробничими процесами [3; 4].

Однією з ключових проблем галузі є ефективність управління мінеральним живленням. Традиційний метод внесення добрив, що домінує з ХХ століття, базується на усередненому нормуванні. Такий підхід передбачає розгляд ґрунту як однорідного середовища, адже не бере до уваги просторову гетерогенність поля та часову мінливість умов вегетації [2]. В результаті виникає ситуація подвійної неефективності.

По-перше, на ділянках з високим рівнем природної родючості або там, де лімітуючим фактором є не елементи живлення, а, наприклад, волога, внесена "середня" норма є надлишковою. Це не призводить до прибавки врожаю, а лише збільшує виробничі витрати та створює ризики вилягання посівів або жирування рослин.

По-друге, на ділянках з низьким забезпеченням елементами живлення "середня" норма виявляється недостатньою для реалізації генетичного потенціалу культури. Відповідно до закону Лібіха, врожайність визначається фактором, що перебуває в найбільшому дефіциті, і його нестачу неможливо компенсувати надлишком інших елементів. Окрім економічних втрат, традиційний підхід породжує значні екологічні ризики. Невикористані рослинами поживні речовини, особливо нітратні форми азоту, легко вимиваються з кореневмісного шару, потрапляючи у ґрунтові води та водойми, що спричиняє їх цвітіння та порушення екосистем.

У відповідь на ці виклики в аграрній науці та практиці сформувався напрям точного землеробства. Його філософія полягає у диференційованому управлінні ресурсам, що передбачає врахування локальних особливостей кожної мікрозони поля [55; 56]. Технологічним інструментом реалізації цього підходу є технології змінного нормування (Variable Rate Technology – VRT), які дозволяють автоматично змінювати норму внесення матеріалів під час руху техніки [9].

Таблиця 1.1

**Порівняння сучасних підходів до управління внесенням добрив та
традиційних методів**

Критерії	Сучасні підходи (Precision Agriculture / Smart Farming)	Традиційні методи
Принцип управління	Диференційований (Site-specific): врахування неоднорідності в межах поля	Усереднений (Uniform rate): одна норма на все поле або масив полів
Інструментарій	Автоматизовані системи, VRT, IoT-сенсори, дрони, супутникові знімки	Ручний відбір проб, паперові агрохімічні карти, загальні довідники
Використання ШІ	Прогнозування врожайності, оптимізація норм (VRT), розпізнавання дефіциту живлення (CV)	Відсутнє, рішення приймаються на основі інтуїції або лінійних формул
Часова дискретність	Моніторинг та корекція у реальному часі (Real-time)	Рішення приймаються 1–2 рази на сезон (перед посівом)
Точність	Висока: локалізація до рівня окремої рослини або мікрозони	Низька: усереднення по великих площах
Переваги	Економія ресурсів, екологічність, максимізація прибутку	Низький поріг входу, простота, відсутність потреби у складному обладнанні
Недоліки	Висока вартість впровадження, потреба у кваліфікованих кадрах, складність інтеграції	Перевитрати добрив, ризик забруднення довкілля, недоотримання врожаю

Джерело: адаптовано та доповнено автором на основі [29; 30]

Впровадження точного землеробства спирається на два основні класи технологій VRT [9; 13], які представлені нижче.

- **Картографічний метод (Map-based VRT).** Використовує попередньо створені карти-завдання, що дозволяє врахувати історичні тренди та дані агрохімічного аналізу. Сучасні дослідження пропонують інтегрувати у цей процес дані БПЛА та GIS-систем для динамічного моніторингу нутрієнтів у ґрунті [11; 12]. Проте метод залишається частково статичним і обмеженим у реакції на миттєві зміни погоди.
- **Сенсорний метод (Sensor-based VRT).** Базується на використанні оптичних сенсорів, які сканують рослини в реальному часі. Дослідження показують ефективність цього методу для азотного підживлення кукурудзи та зернових, де сенсори дозволяють коригувати дозу на основі фактичного стану біомаси [10].

Таблиця 1.2

Класифікація методів реалізації VRT

Характеристика	Картографічний метод (Map-based VRT)	Сенсорний метод (Sensor-based VRT)
Принцип дії	Використовує попередньо створені карти-завдання (prescription maps), завантажені у бортовий комп'ютер. Позиція визначається через GPS/GNSS.	Використовує оптичні сенсори на тракторі/штанзі, що сканують рослини "на льоту". Алгоритм миттєво розраховує потребу в азоті.
Джерела даних	Агрохімічний аналіз ґрунту, карти врожайності минулих років, супутникові індекси, карти електропровідності.	Спектральні індекси (NDVI, NDRE), виміряні активними сенсорами (наприклад, GreenSeeker, Yara N-Sensor).

Характеристика	Картографічний метод (Map-based VRT)	Сенсорний метод (Sensor-based VRT)
Переваги	Дозволяє врахувати історичні дані та комплексні стратегії (Р, К, рН). Точне бюджетування до початку робіт.	Реакція на поточний стан рослин "тут і зараз". Не залежить від застарілих карт. Ідеально для азотного підживлення.
Обмеження	Статичність: карта не враховує зміни погоди, що сталися після її створення. Часомісткість підготовки.	Неможливість врахувати запаси поживних речовин у ґрунті (Р, К). Потребує калібрування під сорт.

Окремим високотехнологічним напрямком є фертигація – метод внесення розчинних добрив разом із поливною водою. Це дозволяє доступно подавати поживні речовини безпосередньо в зону кореневої системи. В умовах краплинного зрошення обсяг зволоженого ґрунту обмежений, тому рослинам складно отримати живлення з периферійних зон. Фертигація розв'язує цю проблему, дозволяючи вносити добрива «мало, але часто», синхронізуючи подачу з фазами росту. Сучасні автоматизовані вузли фертигації використовують датчики рН та електропровідності (ЕС) для контролю концентрації розчину в реальному часі, що мінімізує ризики опіків кореневої системи та підвищує коефіцієнт засвоєння добрив до 90% [16].

У відповідь на зазначені проблеми сформувався підхід точного землеробства та концепція інтегрованого управління поживними речовинами. В основі цього підходу лежить принцип раціонального й збалансованого використання різних джерел живлення рослин. Він передбачає поєднання мінеральних добрив з органічними матеріалами та біологічними препаратами з метою збереження родючості ґрунту і підвищення економічної ефективності

агровиробництва. Застосування такого підходу дає змогу не лише забезпечити рослини необхідними поживними елементами, а й підтримувати довгострокове здоров'я ґрунтового покриву [15].

Аналіз ринку агротехнологій (див. рис. 1.1) свідчить про стрімке зростання інвестицій у цей сектор. За даними Gartner, обсяг світового ринку ІТ в агропромисловості досяг 17,44 млрд доларів США у 2023 році. Ринок пропонує широкий спектр комерційних рішень, таких як платформи супутникового моніторингу (Sentinel Hub, EOS DA), системи мобільної діагностики (Plantix) та експрес-аналізу ґрунту (SoilCares).

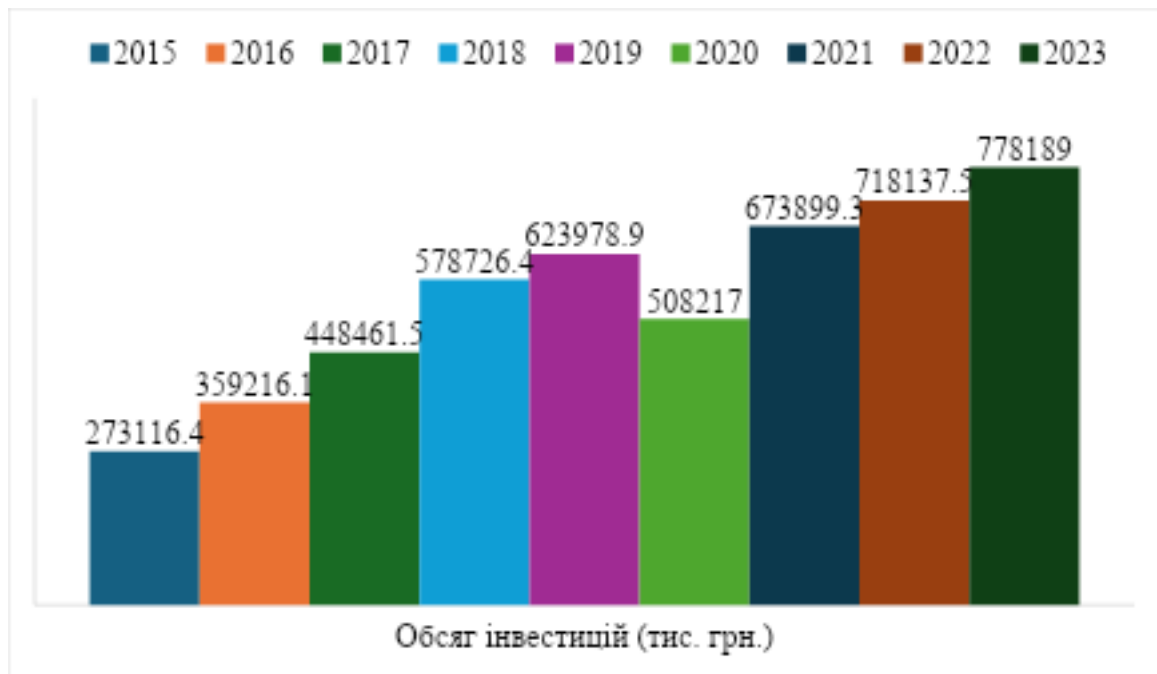


Рис. 1.1. Динаміка капітальних інвестицій у сучасні ІТ у сільському господарстві (2015–2023 рр.)

На сучасному етапі на ринку представлені комерційні рішення, що реалізують окремі елементи концепції точного живлення, зокрема супутниковий моніторинг та експрес-діагностику стану рослин. Огляд найпоширеніших систем подано в таблиці 1.3.

Попри технологічний прогрес, залишається невирішеною методологічна проблема розриву між наявністю даних та якістю управлінських рішень. Більшість наявних систем зосереджується на етапі моніторингу. Вони

відповідають на питання "Який стан поля зараз?". Натомість задача прийняття рішень – "Скільки добрив внести, щоб отримати максимальний прибуток з урахуванням ризиків?", часто вирішується спрощено.

Таблиця 1.3

Комерційні системи інтелектуального управління в агрономії

Сфера застосування	Назва системи / Рішення	Функціональні особливості та роль у системі живлення
Експрес-аналіз тканин	Picketa Systems LENS	Інноваційна система для аналізу рослинної тканини в реальному часі. Дозволяє отримати дані про дефіцит макро- та мікроелементів безпосередньо в полі за лічені хвилини, замінюючи тривалий лабораторний аналіз.
Мобільна діагностика	Plantix	Платформа на базі комп'ютерного зору. Діагностує дефіцит живлення, хвороби та шкідників за фотографіями листя, зробленими смартфоном. Надає рекомендації щодо лікування.
Агрохімічний аналіз	SoilCares / Teralytics	Системи спектрального експрес-аналізу ґрунту. Дозволяють швидко визначити pH, вміст органічної речовини та NPK без використання хімічних реагентів.
Супутниковий моніторинг	Sentinel Hub / EOS DA	Використання супутникових знімків для розрахунку індексів вегетації (NDVI, NDRE), створення карт неоднорідності поля та зонування для VRT.

Сфера застосування	Назва системи / Рішення	Функціональні особливості та роль у системі живлення
Прогнозування врожаю	Ceres Imaging / Yield10	Використання аерофотознімання та ШІ для прогнозування обсягів врожаю та виявлення зон водного/поживного стресу.

Сучасні наукові дослідження намагаються формалізувати цю задачу, використовуючи різні класи математичних моделей:

Послідовні моделі керування (MDP/POMDP). Розглядають процес вирощування як послідовність кроків, де на кожному етапі агент (система) обирає дію (дозу добрив), максимізуючи кумулятивну винагороду. Цей підхід, що використовує методи навчання з підкріпленням, дозволяє враховувати динаміку розвитку рослин, але є складним для реалізації через потребу у величезних обсягах даних для навчання.

Баєсівські ймовірнісні моделі. Дозволяють прямо оцінювати невизначеність. На відміну від детермінованих моделей, які видають одне число (прогноз врожаю), баєсівські підходи моделюють розподіл ймовірностей результату. Це дає змогу оптимізувати рішення не лише за середнім очікуваним прибутком, а й за критерієм мінімізації ризиків.

Системні рамки точного внесення. Інтегрують дані дистанційного зондування та геоінформаційних систем для побудови карт неоднорідності. Оптимізація тут часто зводиться до просторового перерозподілу добрив від гірших зон до кращих, або навпаки, залежно від обраної агрономічної стратегії.

Важливим аспектом, що часто ігнорується в комерційних системах, але підкреслюється в науковій літературі, є верифікація рішень. Традиційна оцінка якості моделей за метриками точності прогнозу є недостатньою для задач управління. Модель може точно прогнозувати врожайність "в середньому", але формувати помилкові рекомендації в умовах аномальних погодних явищ, що

призводить до суттєвих фінансових втрат. Тому виникає потреба в методології наскрізної верифікації, яка перевіряє не лише точність прогнозу, а й економічну ефективність та стійкість генерованих рішень до стохастичних збурень.

Отже, цифровізація засобів моніторингу не супроводжується відповідним розвитком інструментів підтримки прийняття управлінських рішень. Необхідно створити системи, що інтегрують дані сенсорів (зокрема комп'ютерного зору) в моделі стохастичної оптимізації, що і становить предмет цього дослідження.

1.2 Інтелектуальні інформаційні технології в агрономії (підходи, архітектури, аналітичні моделі)

Перейти до концепції інтелектуального сільського господарства неможливо без побудови надійної інформаційної інфраструктури, здатної обробляти великі масиви гетерогенних даних у режимі реального часу. Сучасні агрономічні системи поступово еволюціонують від набору ізольованих інструментів до складних кіберфізичних екосистем. Такі системи поєднують апаратне забезпечення, комунікаційні протоколи та алгоритмічні ядра прийняття рішень. В таблиці 1.4 зазначені основні напрями застосування інтелектуальних технологій в агросекторі.

1.2.1 Багаторівнева архітектура IoT-систем в агрономії. Аналіз джерел свідчить, що еталонна архітектура інтелектуальної системи управління живленням будується на принципах Інтернету речей (IoT) і складається з чотирьох ключових рівнів. Така ієрархія забезпечує інтеграцію вихідних масивів даних для формування інформаційного базису прийняття рішень.

Перший рівень – рівень сприйняття, який можна уявити як «очі» системи. Його завдання полягає у зборі даних та формуванні вхідного вектора для моделей прийняття рішень. Джерела інформації поділяються на контактні та дистанційні сенсори. Контактні сенсори – це IoT-датчики, розміщені безпосередньо в ґрунті, що вимірюють вологість, температуру, електропровідність та рН, забезпечуючи безперервний потік даних про стан кореневого середовища [8; 28]. Дистанційні

сенсори на основі комп'ютерного зору включають камери видимого та мультиспектрального діапазону, встановлені на безпілотних літальних апаратах або наземній техніці. Технології комп'ютерного зору тут не є самоціллю, а виконують функцію потужного сенсора для діагностики фенотипу рослин.

Другий, мережевий рівень забезпечує передачу даних від польових сенсорів до хмарних сервісів. З огляду на особливості агроландшафтів, як-от великі відстані та нерівномірне покриття GSM, визначальну роль відіграють енергоефективні протоколи дальнього радіуса дії (LPWAN), наприклад LoRaWAN та NB-IoT, що дозволяють сенсорам працювати автономно роками [12]. Для передачі «важких» даних, наприклад відеопотоків з дронів, застосовуються локальні мережі Wi-Fi або мобільні мережі 4G/5G.

Таблиця 1.4

Основні напрями застосування інтелектуальних технологій в агросекторі

Напрямок застосування	Технологічна реалізація	Приклад у системі живлення
Моніторинг стану посівів	Супутникові знімки (Sentinel Hub), БПЛА з камерами високої чіткості	Визначення індексу NDVI для побудови карт диференційованого внесення азоту.
Діагностика стресу	Комп'ютерний зір (Computer Vision), згорткові нейромережі (CNN)	Раннє виявлення хлорозу (дефіциту азоту) або хвороб за патернами на листі.
Аналіз ґрунту	ІоТ-датчики, спектрометри (SoilCares)	Автоматична корекція рН та ЕС робочого розчину в системах фертигації.

Напрямок застосування	Технологічна реалізація	Приклад у системі живлення
Прогнозування	Аналіз великих даних (Big Data), рекурентні мережі (LSTM)	Прогноз динаміки вологості ґрунту та врожайності на основі метеоданих.

Джерело: узагальнено автором на основі [2; 14; 16]

Третій рівень – обчислювальний, або Edge-Cloud Continuum передбачає гібридні обчислення. Хмарні обчислення використовуються для навчання складних моделей на історичних даних, зберігання великих масивів інформації та стратегічного планування. Граничні обчислення обробляють дані безпосередньо на борту трактора або дрона, що критично для систем реального часу. Наприклад, у системах точкового обприскування розпізнавання бур'янів і відкриття форсунок має відбуватися за мілісекунди, без затримок.

Четвертий рівень – прикладний, який забезпечує взаємодію агронома з системою. Це мобільні додатки для моніторингу стану посівів, вебпортали для створення карт-завдань та контролери виконавчих механізмів, таких як роботи або системи поливу, що дозволяють реалізувати рішення, рекомендовані аналітичними модулями системи.

Вагомий внесок у розвиток цього напрямку зробили вітчизняні дослідники Національного університету біоресурсів і природокористування України. Зокрема, у роботах Лисенка В.П. та співавторів сформовано науковий напрям інтелектуальних систем керування біотехнологічними об'єктами аграрного призначення, що охоплює розроблення алгоритмів керування робототехнічними комплексами в агровиробництві [5], побудову нейромережевих моделей прогнозування агробіологічних показників на посівах зернових культур [6] та підходи до моделювання продуктивності виробничих процесів методами машинного навчання [7]. Водночас розглянуті в цих та інших роботах архітектури

зосереджені переважно на рівні збору та первинної обробки даних, а також на задачах локального керування технологічними об'єктами, тоді як завдання інтеграції прогнозних моделей з ризик-орієнтованим економічним аналізом для підтримки прийняття рішень в умовах стохастичної невизначеності залишається менш дослідженим, що і визначає фокус подальшого дослідження.

1.2.2 Гібридні прогностичні моделі та роль комп'ютерного зору.

Основою аналітичного блоку системи є алгоритми обробки й аналізу даних, які забезпечують обчислення прогнозних показників на основі вхідних даних. У сучасних дослідженнях спостерігається відхід від простих регресійних моделей до використання методів глибокого навчання, що мають вищу ефективність при роботі з нелінійними біологічними процесами.

Особливе місце в цьому процесі посідає комп'ютерний зір, який трансформувався з інструменту простого моніторингу в потужний засіб діагностики фізіологічного стану рослин. Моделі на базі згорткових нейронних мереж здатні виявляти приховані патерни стресу, аналізуючи текстуру та спектральні характеристики листя. Це дозволяє автоматизувати оцінку забезпеченості рослин азотом та іншими елементами, замінюючи суб'єктивну експертну оцінку на об'єктивні метрики.

Проте використання лише візуальних даних є недостатнім для побудови надійних прогнозів врожайності. Тому провідним трендом є створення гібридних моделей, що об'єднують візуальні ознаки з агрофізичними даними та часовими рядами погодних умов [14].

Сучасний підхід передбачає гібридну архітектуру. Згорткові мережі забезпечують перетворення зображень рослин у вектори ознак для задач фенотипування. Паралельно з цим розвиваються підходи на основі ансамблевих методів градієнтного бустингу та нейромережевих моделей, що забезпечують високу точність регресійного прогнозування цільових показників безпосередньо на основі табличних виробничих, агрокліматичних та агрохімічних даних. Ефективність такого підходу підтверджено зокрема в роботі Лисенка В.П. та

співавторів з моделювання продуктивності цукрового заводу методами машинного навчання, де побудована нейромережева модель забезпечила похибку прогнозу менше 1% [7]. Цей напрям моделювання є основою побудови прогнозного модуля в розробленій інформаційній технології.

1.2.3 Алгоритми прийняття рішень та оптимізації. Ключовою відмінністю інтелектуальних СППР від простих моніторингових систем є наявність блоку прийняття рішень, який визначає оптимальні норми та строки внесення хімікатів. Традиційні підходи до розв'язання цієї задачі базувалися на детермінованих алгоритмах або експертних правилах, які ігнорували невизначеність майбутніх умов. Проте в умовах зміни клімату такий підхід стає неефективний, тож сучасна методологія зміщується в бік стохастичної оптимізації, де врожайність та економічний результат розглядаються не як фіксовані величини, а як випадкові змінні з певним розподілом ймовірностей.

Задача управління живленням у такій постановці формалізується як пошук оптимальної стратегії, що максимізує цільову функцію при заданих обмеженнях. На відміну від задач класичної оптимізації, тут критично важливо врахувати ризики. Оптимальна доза добрив, розрахована для "середнього" року, може призвести до значних збитків у випадку посухи або інших аномалій. Тому передові системи інтегрують у цільову функцію метрики ризику, такі як CVaR (Conditional Value-at-Risk), що дозволяє знаходити рішення, стійкі до найгірших сценаріїв розвитку подій. Це означає перехід від максимізації врожаю за будь-яку ціну до максимізації економічної ефективності з контрольованим рівнем ризику [21].

Реалізація таких алгоритмів передбачає використання методів навчання з підкріпленням або сценарного моделювання, у межах яких система аналізує множину можливих кліматичних сценаріїв та формує стійку стратегію управління. Важливою вимогою до таких систем є наскрізна верифікація рішень, що передбачає перевірку не лише точності проміжного прогнозу врожайності, а й економічної коректності кінцевої рекомендації. Це вимагає створення

спеціальних протоколів тестування, що оцінюють поведінку оптимізатора в широкому спектрі синтетичних та історичних сценаріїв, гарантуючи, що запропоноване рішення не призведе до деградації економічних показників порівняно з базовими стратегіями.

Комплексне застосування засобів збору даних та моделей оптимізації дозволяє створювати системи адаптивного управління ризиками в умовах невизначеності.

1.3. Проблеми адаптації моделей прийняття рішень до змінних природно-кліматичних факторів

Основною проблемою впровадження оптимізаційних систем є висока мінливість умов довкілля. На відміну від промислових об'єктів зі стабільними та контрольованими умовами, агроecosистеми є відкритими стохастичними системами, на які впливає низка непередбачуваних чинників. Традиційні моделі машинного навчання, навчені на історичних даних, часто демонструють різке зниження ефективності при розгортанні в реальних умовах через низку проблем, що вимагають перегляду принципів їх побудови й оцінювання.

1.3.1 Дрейф концепції та нестационарність даних. У класичному машинному навчанні є фундаментальне припущення, що навчальна вибірка та дані, на яких модель працюватиме в майбутньому, мають однаковий статистичний розподіл. В агрономії це припущення майже завжди порушується через явище дрейфу концепції [19]. Це значить, що статистичні властивості цільової змінної, наприклад, залежність врожайності від дози азоту, змінюються з часом під впливом глобальних кліматичних зрушень.

Модель, навчена на історичних даних, може рекомендувати заходи, ефективні для вологих умов минулого. Однак в умовах сучасного потепління ці рекомендації стають помилковими або навіть шкідливими [17]. Фізіологічна реакція рослин на добрива при дефіциті вологи кардинально змінюється,

перетворюючи поживні речовини з фактора росту на фактор сольового стресу, що робить історичні залежності "вхід-вихід" неактуальними [18].

Окрім часової нестаціонарності, значною проблемою є просторова варіабельність та потреба в адаптації домену. Модель, що показує високу точність на полях одного регіону, часто виявляється непридатною для іншого через відмінності в типах ґрунтів, мікрокліматі та сортах культур. Збір великих розмічених наборів даних для кожного нового поля є економічно недоцільним, що актуалізує проблему перенесення знань. Без ефективних механізмів адаптації, які б дали змогу "доналаштувати" глобальні моделі на локальних даних з мінімальними витратами, масштабування інтелектуальних систем стає неможливим [20].

1.3.2. Обмеженість детермінованих підходів та ігнорування ризиків. Ще однією системною проблемою є домінування детермінованих підходів до оптимізації, які ігнорують стохастичну природу аграрних ризиків. Більшість наявних систем підтримки рішень розраховують рекомендації, виходячи з прогнозу врожайності для "середніх" або "найімовірніших" погодних умов. Такий підхід є ризиконейтральним і не враховує асиметрію наслідків прийнятих рішень. В агровиробництві втрати від несприятливих подій (наприклад, посухи) часто є фатальними для економіки господарства, тоді як додатковий прибуток у сприятливі роки не компенсує ризику банкрутства. Норма добрив, розрахована для кліматичної норми, у посушливий рік може стати токсичною для рослин або економічно збитковою через низький коефіцієнт засвоєння, що підкреслює вразливість детермінованих стратегій.

Сучасна теорія управління вказує на потребу в переходу до методів надійної оптимізації та стохастичного програмування, що явно інтегрують міри ризику, такі як умовна вартість під ризиком (CVaR) [21]. Мета такої системи полягає не просто в максимізації математичного сподівання прибутку, а й у формуванні стратегії, що залишається стійкою в широкому діапазоні можливих сценаріїв та мінімізує ймовірність критичних збитків ("хвостових" ризиків). Проте більшість

комерційних та наукових розробок досі ігнорують цей аспект, пропонуючи рішення, які є математично оптимальними лише на папері, але крихкими в реальному житті [22].

1.3.3. Прийняття рішень в умовах невизначеності. Менш дослідженим, проте істотним аспектом є верифікація якості управлінських рішень. У науковій літературі є розрив між верифікацією прогнозової складової та верифікацією кінцевого результату. Дослідники використовують метрики точності прогнозу, такі як середньоквадратична похибка (RMSE) або коефіцієнт детермінації (R^2), для оцінки моделей. Проте висока точність прогнозу врожайності не гарантує економічну ефективність рекомендованої дози добрив [21]. Невелика систематична помилка в оцінці функції відгуку може значно змістити точку оптимуму і, як наслідок, призвести до суттєвих фінансових втрат.

Ситуація ускладнюється відсутністю уніфікованих протоколів наскрізної верифікації, які б оцінювали весь ланцюжок "дані – прогноз – оптимізація – рішення". Сучасні підходи часто не перевіряють стійкість рішень до зміни вхідних даних та не порівнюють їх з базовими (наївними) стратегіями, що створює ризик впровадження складних, але практично неефективних систем. Об'єктивні труднощі проведення польових експериментів, які вимагають багаторічних спостережень, призводять до домінування результатів, отриманих у симуляційному середовищі, що породжує проблему розриву між моделюванням та реальністю. Розв'язання цієї проблеми вимагає розробки нових методологій, які поєднують сценарний аналіз, історичну реконструкцію та економічне моделювання для підтвердження надійності пропонованих технологій.

Науковий пошук зміщується від розробки нових архітектур до методології їх адаптації та верифікації. Критичним завданням стає створення доказової бази, що підтверджує економічну ефективність та надійність інтелектуальних рішень у нестационарному середовищі.

1.3.4 Сучасні підходи до математичної формалізації задачі оптимізації внесення добрив. У сучасних дослідженнях з точного землеробства формалізація

задачі оптимізації внесення мінеральних добрив у стохастичних умовах найчастіше ґрунтується на чотирьох взаємодоповнювальних класах моделей: послідовні моделі керування (MDP/POMDP) з підкріпленням, де оптимальна норма внесення є результатом політики прийняття рішень у часі; прогностичні ML-моделі відгуку врожайності з подальшою оптимізацією економічного критерію; ймовірнісні (байєсівські/PDL) моделі, які безпосередньо оцінюють невизначеність і дозволяють оптимізувати не лише очікуваний ефект, а й ризик; концептуальні моделі диференційованого внесення (remote sensing + GIS + VRT), де оптимізація керованих змінних пов'язана з просторовою неоднорідністю поля та адаптивною актуалізацією знань.

Перший напрям формалізації розглядає управління живленням як задачу динамічного керування. Зокрема, у роботі [23] застосовано методи глибокого навчання з підкріпленням у поєднанні з імітаційним моделюванням агроєкосистем. Важливість цієї формалізації полягає в переході від фіксованого сезонного нормативу до послідовної оптимізації: оптимальна дія визначається як елемент множини допустимих дій на кожному кроці часу, а ціль – максимізація сумарного економічного ефекту, що відображає компроміс між продуктивністю та витратами або втратами елементів живлення. Обґрунтовано, що вектор норми внесення добрив є не константою, а траєкторією керування, а стохастичність (погодні умови, реакція культури) природно інкапсулюється у динаміці середовища та варіативності епізодів навчання, навіть коли явна аналітична функція врожайності від N недоступна або є істотно нелінійною [23].

Цей підхід розширено у дослідженні [28], к частково спостережуваний марковський процес прийняття рішень (POMDP): середовище описується кортежем $\langle S, A, O, T, R \rangle$, де S – простір прихованих станів агроєкосистеми, A – множина управлінських дій (зокрема, норми внесення), O – спостереження, T – перехідна динаміка, R – винагорода. Ключовим у такій формалізації є те, що оптимізація здійснюється за комплексним показником, що узгоджує агрономічний результат, операційні витрати та екологічні наслідки: автори акцентують увагу на

балансі між доходами від урожаю та екологічними наслідками за надмірне використання ресурсів, вимивання нітратів і викиди. У межах зазначеної роботи важливим є врахування невизначеності через стохастичні сценарії клімату та аналіз розподілів очікуваних результатів за допомогою прогнозних інтервалів, що дозволяє оцінювати не тільки середній ефект політики, а й її стійкість до погодних збурень. Доведено, що оптимальні політики суттєво підвищують винагороду в екстремальних сценаріях (наприклад, при критичному зменшенні опадів), однак при цьому здатні зміщувати компроміс у бік інших екологічних метрик, що наочно ілюструє вартісну оцінку компромісу вибраної структури R .

На відміну від підходів на основі навчання з підкріпленням (RL), які трактують оптимізацію як пошук політики, класичні статистичні постановки явно вводять невизначеність у вигляді розподілів параметрів. Так, Міа та колеги формалізують вплив норм удобрення на врожайність у схемі польових експериментів (on-farm experimentation) із просторовою неоднорідністю та застосовують моделі на основі гауссівських процесів (Gaussian process model), які дозволяють врахувати просторові варіації врожайності як джерело стохастики та похибок вимірювання [26]. Далі критерій оптимізації набуває економічного змісту через оцінку нетто-доходу на гектар для різних сценаріїв норм внесення, причому розрахунок виконується на кожній байєсівській ітерації, що дає розподіл доходу, а не точкову оцінку [26]. Це безпосередньо пов'язує оптимізацію параметрів із ризик-орієнтованим підходом: автори оцінюють, яка частка польових випадків забезпечує позитивний дохід за фіксованих рівнів імовірності (наприклад 95/90/80%), і будують просторові карти ймовірностей з урахуванням варіацій цін на продукцію та добрива [26]. Вимірюваними результатами виступають не тільки зміни врожаю, але й імовірність позитивного фінансового ефекту, що методологічно важливо для стохастичних умов: оптимальною стає стратегія, яка є вигідною не усереднено, а з прийнятною гарантією [26].

Альтернативним підходом до урахування ймовірнісних факторів – це навчання моделей, які одразу повертають розподіл або інтервал прогнозу. У межах

цієї логіки Wang разом із колегами демонструють ефективність імовірнісного глибокого навчання (Probabilistic Deep Learning, PDL), де вихід моделі параметризує логнормальний розподіл $\ln(X) \sim N(\mu, \sigma^2)$, а навчання здійснюється мінімізацією від'ємної логарифмічної правдоподібності $\min \sum_{i=1}^n \log P(y_i | x_i; \theta)$ [28]. Кість такої моделі оцінюється не тільки за RMSE, а й за метриками покриття інтервалу, зокрема ймовірністю покриття прогнозного інтервалу (PICP), що визначається як частка спостережень, які потрапили у межі інтервалу [28]. Отримане значення PICP=0.937 для 95% інтервалів є кількісним підтвердженням того, що модель утримуватиме невизначеність у близьких до номінальних межах [28]. Наукова цінність для дослідження полягає в тому, що оптимізація керованої змінної може бути побудована на консервативних або робастних оцінках, які підвищують стійкість рішень у змінному кліматі та за умов неповноти даних.

На противагу політикам у часі, значна частина робіт розглядає інтелектуальну підтримку внесення добрив як задачу рекомендацій або прогнозу, де формалізація спирається на точність моделей та продукційні правила (rule-based). У статті Musanase та колег система рекомендацій будується як комбінація двох підмоделей: нейромережевої моделі для рекомендації культур за параметрами росту та ґрунтовими показниками і моделі на основі правил, що видає поради щодо удобрення за попередньо сформованими таблицями [24]. Автори декларують точність прогнозової частини на рівні 97%, що інтерпретується як висока придатність для прийняття управлінських рішень у точному землеробстві [24]. Зазначений підхід є прикладом формалізації, де критерій оптимальності задається не через явну цільову функцію, а через відповідність рекомендацій еталонним класам: така постановка добре масштабується, але має обмежені можливості щодо врахування стохастичності як ризику, якщо не додати ймовірнісні оцінки або економічні функції корисності [24].

Особливості застосування методів машинного навчання для оцінювання точності прогнозування розглянуто в роботі Hasan та ін. Автори порівнюють

традиційні ансамблеві алгоритми та показують доцільність вибору моделі за метриками похибки MAE/MSE/RMSE та коефіцієнтом детермінації у задачах, чутливих до погодної невизначеності [25]. Дослідження підтверджує, що запропонована модель Kernel Ridge Regression (KRR) мінімізує помилки (зокрема, відносні показники MAE і RMSE) та зберігає високу якість на даних з аномальними роками (наприклад, періоди, коли повені або стреси порушують звичну динаміку врожаю) [25]. У контексті нашого дослідження цей результат засвідчує, що формалізація оптимізації може спиратися на стабільний прогностичний блок як функцію відгуку в умовах стохастичних кліматичних траєкторій: тоді оптимізація норми внесення стає наступним етапом на базі надійного прогнозу [25].

Показовою для інтелектуальних ІТ у внесенні добрив є напрям досліджень, де оптимізація інтегрується з дистанційним моніторингом та просторовою аналітикою. Chen та колеги пропонують формалізацію «точного внесення» як замкнений цикл «моніторинг → моделювання → рішення → оновлення», де рішення генеруються на основі моделі злиття даних (мультиспектральна зйомка з БПЛА та наземні сенсори) та GIS-аналізу [11]. У роботі явно задаються математичні компоненти системи: модель спектральної відповіді $R(\lambda) = \sum_{i=1}^n w_i S_i(\lambda) + \varepsilon(\lambda)$ та формула зваженого злиття джерел $F(x, y) = \alpha H(x, y) + \beta M(x, y) + \gamma G(x, y)$ з оптимізацією ваг за критерієм помилки реконструкції [11]. Вимірювані результати включають істотні покращення точності оцінки вмісту N, P, K (таблично наведені RMSE і R^2 із 95% CI), а також практичні ефекти: зменшення внесення добрив на 18–27% при підвищенні врожаю на 4–11% та підвищенні ефективності засвоєння добрив на 36–48% [11]. У межах нашого напрямку ця формалізація демонструє, що оптимізація «доз» у стохастичних умовах часто починається з оптимізації інформації про стан (оцінка NPK у просторі й часі) та її невизначеності; і лише після цього доза стає контрольованою змінною, яку можна оптимізувати економічно чи за корисністю [11].

Паралельно з методами ML та RL існують системи, де формалізація зосереджена на знання-орієнтованих структурах і відтворюваності експертної логіки. León Chilito разом із колегами розробляють систему рекомендацій добрив за методологією CRISP-DM і будують модель міркувань на основі прецедентів (Case-Based Reasoning, CBR), яка зіставляє поточну ситуацію господарства з базою випадків і видає рекомендовані норми та частоти внесення N, P, K [30]. Вимірювані результати подано як експертну оцінку ефективності рекомендацій, а також як конкретні матриці норм залежно від щільності посадки та стану ґрунту [30]. Методологічна цінність цієї постановки полягає в тому, що стохастика інкапсулюється в різноманітні кейсів і змінності контекстів, але кількісно не оцінюється як ризик; отже, CBR-підхід доцільно комбінувати з ML-блоком прогнозу та економічним модулем оптимізації, якщо потрібно перейти від узгодженості з експертами до максимізації очікуваного прибутку під імовірнісними обмеженнями [30].

Окремий клас робіт спрямований на інтерпретованість рекомендацій, де формалізація задачі задається як класифікація типу або категорії добрива з пояснюваністю рішень. Venkateswara та колеги демонструють інтерпретовані deep learning моделі для незалежних рекомендацій цільової культури та типу добрива, оцінюючи їх через метрики якості (accuracy, precision, recall, F1, AUC, MCC) та крос-валідацію [27]. Вимірювані результати показують середню точність приблизно у 95.24% при 5-fold CV, високі значення AUC і MCC $\approx 94.74\%$, а також наявність механізмів інтерпретації на кшталт SHAP [27]. Це свідчить про те, що формалізація процесу видачі рекомендацій може бути відтворюваною й пояснюваною, але для задачі оптимізації доз у стохастичних умовах цього недостатньо без переходу від дискретної рекомендації до безперервної керованої змінної (норма N/P/K) і без визначення явної цільової функції (прибуток, ризик, екологічний штраф). Проте інструменти пояснюваності є вкрай корисними як частина інтелектуальної IT-системи для обґрунтування вибору моделі та впливу факторів на кінцеве рішення [27].

Узагальнення наявних оглядових синтезів дозволяє уточнити, як саме в літературі систематизується формалізація оптимізації через технологічні підсистеми. У PRISMA-огляді Cai та колеги підкреслюють, що оптимізація хімічних добрив у точному землеробстві найчастіше реалізується через зв'язку дистанційного зондування, ГІС-картування та технологій диференційованого внесення (VRT), де кінцевим результатом алгоритму є формування карти-завдання і зональна норма внесення, а стохастичність – це невизначеність спостережень, часові коливання та неоднорідність поля [2]. Це формує концептуальний базис: оптимізація норми є частиною загальної задачі оптимізації процесу внесення (коли, де, скільки), а інтелектуальні технології слугують механізмом перетворення даних на керовані параметри внесення в умовах невизначеності [2].

Аналіз наведених джерел свідчить про можливість формалізації задачі оптимізації як вибору керованого вектора доз (або політики) під випадковими збуреннями (погода, ціни, спостереження стану, просторові флуктуації), де критерій оптимальності визначається або максимізацією очікуваної корисності чи прибутку з екологічними штрафами (парадигма RL/POMDP) [28], або максимізацією ймовірності позитивного фінансового результату чи мінімізацією ризику через байєсівський розподіл ефектів (парадигма GP) [26], або оптимізацією у просторі через високоточну оцінку NPK та карти-завдання (prescription maps) (парадигма UAV+GIS+ML) [11], або ж задачею рекомендацій (ML/CBR), де формалізація зводиться до класів рекомендацій та їхньої валідності [24, 30] і може бути доповнена інтерпретованістю як важливою вимогою до системи [27].

1.3.5 Сучасні підходи до управління внесенням мінеральних добрив: класифікація і порівняльний аналіз. Серед сучасних наукових підходів до формування рекомендацій з нормування мінеральних добрив доцільно виокремити чотири репрезентативні класи методик, що відрізняються вихідними принципами формування рекомендованої дози: традиційні схеми фіксованих

рекомендацій, сенсорно-керовані підходи точного внесення, моделі прийняття рішень на основі машинного навчання з оптимізаційним пошуком та партисипативні підходи із залученням виробників.

Традиційну схему фіксованих рекомендацій (Recommended Dose of Fertilizer, RDF) представлено в роботі Sairam et al. [50], виконаній на дослідних ділянках Паралахемунді (штат Одіша, Індія) для кукурудзи. Норма внесення формується на основі агрохімічного балансу ґрунту й узагальнених рекомендацій для регіону без урахування погодної варіативності конкретного сезону; отримана базова норма становить 120,0-26,2-49,8 кг/га за діючою речовиною N, P, K відповідно.

Сенсорно-кероване точне внесення розглянуто в роботі Sujatha et al. [100], де норма азоту для кукурудзи на дослідних ділянках Педдапурам (штат Андхра-Прадеш, Індія) коригується за оптичним індикатором стану посіву в реальному часі. На відміну від RDF-підходу, цей метод реагує на фактичний фізіологічний стан рослин, проте не передбачає явного моделювання погодного чи цінового ризику; сформована референтна норма становить 164,5-26,2-41,5 кг/га.

Модель прийняття рішень на основі машинного навчання з оптимізаційним пошуком представлено в роботі Gao et al. [101], де для умов Венсу (Сіньцзян-Уйгурський автономний район, Китай) норму внесення для кукурудзи визначено за допомогою алгоритму зозулі (Cuckoo Search Algorithm) на основі ML-моделі відгуку врожайності; отримана рекомендація становить 178,6-33,0-55,7 кг/га. Цей підхід методологічно близький до напрямку, розглянутого в підрозділі 1.3.4 (прогностичні ML-моделі з подальшою оптимізацією економічного критерію), проте обмежується точковою оптимізацією без явного представлення розподілу можливих результатів.

Партисипативний підхід до управління живленням рослин із залученням виробників (Farmer-Participation Nutrient Management, FPNM) застосовано в роботі Ren et al. [102] для озимої пшениці в провінції Хенань (Китай). Норма NPK на рівні 223,9-49,5-45,5 кг/га формується на основі узгодження агрономічних

рекомендацій з практичним досвідом і ресурсними можливостями господарств, що відрізняє цей підхід від суто алгоритмічних методик.

Окремої уваги потребує підхід, запропонований у роботі [51], де для визначення економічно оптимальної норми азоту для конкретних ділянок (site-specific economic optimal nitrogen rate, EONR) застосовано моделі Generalized Additive Models та випадкового лісу, навчені на даних точних внутрішньогосподарських експериментів (on-farm precision experimentation). На відміну від чотирьох розглянутих вище підходів, цей метод орієнтований виключно на оптимізацію норми азоту без визначення повного вектора NPK, що обмежує можливості його прямого зіставлення з підходами, які формують комплексну рекомендацію за трьома елементами живлення.

Узагальнюючий аналіз чотирьох основних розглянутих підходів свідчить про їхню спільну методологічну властивість: кожен із них формує єдину рекомендовану дозу за певним критерієм: агрохімічним балансом, фізіологічним станом посіву, алгоритмічною оптимізацією або експертно-практичним консенсусом, але жоден не супроводжує цю рекомендацію кількісною характеристикою фінансового ризику в умовах погодної та цінової невизначеності. Аналіз ризику, якщо присутній, обмежується якісним обговоренням стійкості підходу, але не входить до формального критерію формування рекомендації. Це визначає методологічний розрив, подоланню якого присвячено подальші розділи дисертаційної роботи: перехід від точкової рекомендації до структурованого простору ризик-економічних альтернатив, що відображає компроміс між очікуваним прибутком та стійкістю рішення в несприятливих сценаріях. Детальний кількісний аналіз цього переходу з використанням обраних чотирьох підходів як орієнтирів для обчислювальної апробації наведено в розділі 4.

1.4. Узагальнення теоретичних положень та постановка наукового завдання

Результати аналізу літературних джерел дозволяють констатувати наявність системного методологічного розриву в сучасних інформаційних технологіях підтримки прийняття рішень (СППР) у сфері нормування мінеральних добрив. Розвиток методів машинного навчання, зокрема ансамблевих алгоритмів градієнтного бустингу, забезпечив побудову нелінійних прогнозних моделей врожайності з високою точністю на глобальних і регіональних наборах даних. Попри це, більшість наявних підходів до управління внесенням ресурсів залишаються детермінованими: рекомендаційні системи формують точкові оцінки оптимальної норми внесення без кількісної характеристики розподілу можливих економічних наслідків в умовах погодної та цінової невизначеності. Виявлено, що типові класи СППР — нормативні схеми, сенсорно-керовані підходи на основі вимірювання азотного статусу рослин, регіональні ML-оптимізаційні моделі та інтерактивні системи опитування фермерів — мають спільне обмеження: усі вони генерують фіксовану рекомендацію без явного моделювання ризику. Оцінювання ризиків у таких системах має переважно дескриптивний характер і не інтегроване у формальний оптимізаційний критерій.

Отже, існує суттєва суперечність між високою точністю сучасних прогнозних моделей врожайності та обмеженою здатністю трансформувати ці прогнози в економічно обґрунтовані рішення з кількісним оцінюванням ризику. Крім того, розглянуті профільні дослідження в галузі агрономічного прогнозування здебільшого зосереджені на метриках якості самих моделей як самостійному результаті. При цьому поза увагою залишаються питання інтеграції математичного апарату в програмну платформу, придатну для практичного використання особою, яка приймає рішення (ОПР).

Розв'язання цієї суперечності полягає в розробленні гібридних інтелектуальних інформаційних технологій, що поєднують нелінійне

прогнозування методами машинного навчання з ризик-орієнтованим економічним аналізом на сценарному просторі погодних реалізацій. Такий підхід дозволяє перейти від точкових рекомендацій до надання ОПР повного простору ризико-економічних компромісів між очікуваним прибутком та стійкістю системи у несприятливих сценаріях.

З огляду на це, подальші етапи дисертаційного дослідження спрямовані на створення методологічного, математичного та програмно-інженерного забезпечення такої інформаційної технології. Для досягнення визначеної мети передбачено вирішення таких завдань: розроблення інтегрованої методики формування простору ознак на основі відкритих джерел даних; побудова нелінійної прогнозової моделі врожайності з обґрунтуванням внеску структурних регресорів; формалізація задачі оптимізації внесення добрив за критерієм прибутковості з урахуванням ризику на основі поєднання підходу Mean-Variance та метрики Conditional Value-at-Risk. Практична цінність роботи полягає у програмній реалізації гібридної інформаційної технології у двокомпонентній архітектурі з нативним інференсом серіалізованої моделі. Заверхальним етапом є експериментальна верифікація розробленої технології шляхом компаративного аналізу з альтернативними підходами на релевантних географічних локаціях.

Висновки до розділу 1

У розділі здійснено теоретико-методологічний аналіз проблеми створення інтелектуальних систем підтримки прийняття рішень в агрономії. Основні результати полягають у наступному:

Встановлено, що перехід до концепції «Сільське господарство 4.0» вимагає зміни парадигми управління: від використання усереднених норм добрив до адаптивних стратегій точного землеробства. Показано, що ключовим бар'єром ефективності є не дефіцит технічних засобів чи даних моніторингу, а відсутність адекватних алгоритмів, здатних генерувати оптимальні рішення в умовах високої природно-кліматичної невизначеності.

Обґрунтовано доцільність застосування гібридних архітектур систем підтримки прийняття рішень, що поєднують методи машинного навчання з кількісним моделюванням ризику. Показано, що ключовою методологічною властивістю таких архітектур є розділення обчислювально витратного етапу навчання прогнозовної моделі від етапу її експлуатації, що сприяє масштабованості рішень у виробничих умовах. Визначено, що розглянуті в підрозділі 1.2.1 архітектурні рішення вітчизняних та закордонних дослідників зосереджені переважно на рівні збору й первинної обробки агробіологічних даних, тоді як інтеграція прогнозних моделей з ризик-орієнтованим економічним аналізом залишається менш дослідженою.

Аргументовано необхідність переходу від детермінованих підходів до методів стохастичної оптимізації. Аналіз показав, що ігнорування ймовірнісної природи аграрних процесів призводить до рішень, нестійких до погодних аномалій. Використання ризикоорієнтованих критеріїв дає змогу математично формалізувати стратегію, спрямовану на пошук компромісу між максимізацією прибутку та мінімізацією можливих збитків.

Виявлено методологічний розрив у наявних дослідженнях між оцінкою точності прогнозовної моделі й оцінкою якості управлінського рішення, що з неї випливає. Обґрунтовано необхідність переходу від точкових рекомендацій до структурованого підходу, який дозволяє оцінювати економічну ефективність та стійкість запропонованого рішення до погодних і цінових сценаріїв ще на етапі моделювання, без проведення реального польового експерименту.

Проведений аналіз створює теоретичні передумови для подальшої розробки математичних моделей та програмних засобів, представлених у наступних розділах роботи.

РОЗДІЛ 2

МАТЕМАТИЧНА МОДЕЛЬ ГІБРИДНОЇ СИСТЕМИ ОПТИМІЗАЦІЇ ДОЗ ДОБРІВ

2.1. Формалізація задачі оптимізації доз мінеральних добрив у стохастичних умовах

2.1.1 Стохастична модель врожайності як функція керованих агрохімічних рішень та природно-кліматичних факторів. Нехай на ймовірнісному просторі (Ω, F, P) задано випадковий вектор сезонних погодних умов W , який для кожної країни c реалізується як:

$$W_c = (R_c, GDD_c, SDD_c, \bar{T}_c), \quad (2.1)$$

де R – сумарні опади вегетаційного періоду (*TotalRainGrowingSeason*), GDD – сума ефективних температур, SDD – показник температурного стресу, $\bar{T}$ – середня температура сезону (*AvgTempGrowingSeason*). Керованим рішенням є вектор внесення діючої речовини $u = (N, P, K)$, що безпосередньо відповідає кількості нітратів, фосфатів, та калію відповідно. Важливо, що керування ведеться саме у просторі діючої речовини, тобто якщо застосовується певний продукт добрива з часткою діючого елемента $\rho_N \in (0, 1]$, то маса товарного добрива m_N пов'язана з керованою змінною як $N = \rho_N m_N$, а отже економічна вартість одного кілограма діючої речовини коректно порівнюється між різними продуктами через приведення до діючої речовини (аналогічно для P, K):

$$c_N^{(dr)} = c_N^{(prod)} / \rho_N \quad (2.2)$$

Така нормалізація також узгоджується з необхідністю розділяти N, P, K за законом мінімуму Лібіха: на практиці врожайність може обмежуватися дефіцитом одного елемента, навіть якщо інші внесені надлишково. Тому агрегування у сумарну масу добрив математично руйнує структуру задачі.

Врожайність Y розглядається як випадкова величина. З огляду на емпіричну асиметрію врожайності та потребу стабілізувати дисперсію, модель будується для

логарифмічної врожайності $Z = \log Y$. Таке перетворення має математичне обґрунтування: за умови мультиплікативного характеру похибок біологічних процесів логічною є формула:

$$Y = \tilde{Y} \cdot \eta, \eta > 0, \quad (2.3)$$

де $\tilde{Y}$ – детермінована компонента (очікувана врожайність за фіксованих умов), а η – випадковий мультиплікативний шум. Тоді:

$$\log Y = \log \tilde{Y} + \log \eta, \quad (2.4)$$

тобто випадкова складова стає адитивною, що суттєво спрощує оцінювання та контроль похибок. Позначимо

$$Z_c = \log Y_c. \quad (2.5)$$

Тоді загальна структурна модель відгуку записується як:

$$Z_c = g_\theta(x_c(u)) + \varepsilon_c, \quad (2.6)$$

де $g_\theta(\cdot)$ – нелінійна функція відгуку (що в реалізації буде відтворена ML-моделлю), ε_c – випадкова похибка з $E[\varepsilon_c | x_c] = 0$, а $x_c(u)$ – вектор ознак:

$$x_c(u) = (N, P, K, R_c, GDD_c, SDD_c, CTL_c), \quad (2.7)$$

де CTL_c – змінна CountryTechLevel, яка відображає середній агрономічний потенціал країни і визначається як середнє по розширеному переліку продовольчих культур S співвідношення врожайності країни до середньосвітової врожайності відповідної культури (2.8). Перелік S сформовано на основі даних FAOSTAT про обсяги світового виробництва і включає культури з широким географічним ареалом вирощування, а також обмежений набір теплолюбних культур для коректного покриття країн тропічного агрокліматичного поясу.

$$CTL_c = \left(\frac{1}{|S|} \right) \sum_{s \in S} \left(\frac{\bar{Y}_{c,s}}{\bar{Y}_{world,s}} \right), \quad (2.8)$$

де S – перелік обраних продовольчих культур, $\bar{Y}_{c,s}$ – середня врожайність культури s у країні c за весь період спостережень, $\bar{Y}_{world,s}$ – середньосвітова врожайність культури s за той самий період.

З математичної точки зору CTL_c є оцінкою контекстного зсуву для конкретної країни у просторі врожайності, тобто наближує латентну змінну технологій, менеджменту та ресурсної забезпеченості. Якщо переписати модель у вигляді частково-ієрархічної декомпозиції

$$Z_c = \alpha(CTL_c) + h_\theta(N, P, K, R, GDD, SDD) + \varepsilon_c, \quad (2.9)$$

то CTL_c виступає як структурний регресор, що зменшує зміщення, пов'язане з неврахованими довгостроковими факторами, які не входять у сезонні погодні змінні.

З огляду на закон мінімуму Лібіха, вектор (N, P, K) не може бути зведений до однієї агрегованої змінної. У найпростішому теоретичному наближенні лімітуючого фактора врожайність може бути описана як:

$$Y = Y_{\max}(W) \cdot \min\{\phi_N(N), \phi_P(P), \phi_K(K)\}, \quad (2.10)$$

де $Y_{\max}(W)$ – зумовлений погодою потенціал врожаю, а $\phi_i(\cdot) \in [0, 1]$ – функції забезпеченості елементом живлення (наприклад, насичувальні криві). Якщо ж спробувати замінити (N, P, K) на суму $S = N + P + K$, то мін-структура зникає: не існує такої $\psi(S)$, що для всіх N, P, K виконувалося би:

$$\min\{\phi_N(N), \phi_P(P), \phi_K(K)\} = \psi(N + P + K). \quad (2.11)$$

Це означає, що агрегування спотворює постановку: дві комбінації (N, P, K) з однаковою сумою S можуть давати принципово різний результат залежно від того, який елемент є лімітуючим.

Задача оптимізації формулюється через економічний критерій, оскільки в агровиробництві кероване рішення оцінюється не тільки врожаєм, а чистим прибутком.

Нехай p – ціна продукції (ціна культури), c_N, c_P, c_K – питомі вартості одного кілограма діючої речовини відповідних елементів, C_0 – фіксовані операційні витрати незалежні від дози. Тоді прибуток на гектар при реалізації сезону s за рішення u дорівнює

$$P_s(u) = p Y_s(u, W_s) - (c_N N + c_P P + c_K K) - C_0. \quad (2.12)$$

Оскільки в моделі прогнозується $Z = \log Y$, то $Y = \exp(Z)$, і прогнозована врожайність через функцію g_θ набуває вигляду:

$$Y = \exp(\hat{Z}_c(u)) = \exp(g_\theta(x_c(u))). \quad (2.13)$$

За наявності адитивної похибки у лог-просторі $\varepsilon \sim N(0, \sigma^2)$ математичне сподівання врожайності в оригінальному просторі дорівнює:

$$E[Y | x] = E[\exp(Z) | x] = \exp\left(E[Z | x] + \frac{1}{2}\sigma^2\right), \quad (2.14)$$

що є прямим наслідком логнормального розподілу. Звідси впливає механізм визначення поправки на зміщення при переході з лог-простору в початковий простір [57; 58]: якщо $\hat{Z} = g_\theta(x)$ апроксимує $E[Z|x]$, то незміщена оцінка умовного середнього врожаю має вигляд

$$\hat{Y} = \exp(\hat{Z} + \frac{1}{2}\hat{\sigma}^2), \quad (2.15)$$

де $\hat{\sigma}^2$ – оцінка дисперсії залишків на валідації. Це має критичне значення для оптимізації за прибутковістю, бо прибуток лінійний за Y і помилка у відновленні масштабу безпосередньо зміщує оптимум.

Стохастичність умов проявляється через W_c і (ширше) через ε_c . З огляду на це, математично обґрунтованою є постановка оптимізації не за реалізацією Π , а за її функціоналом ризику. Базова (ризик-нейтральна) постановка – максимізація очікуваного прибутку:

$$u^* = \arg \max_{u \in \Omega} E[\Pi_c(u)], \quad (2.16)$$

де Ω – допустима область норм внесення. Оскільки рішення приймається до реалізації погоди, математичне сподівання береться за розподілом W_c та інших джерел невизначеності. У програмній реалізації задача розв'язується методами Монте-Карло. Генерується множина сценаріїв $\{W^{(s)}\}_{s=1}^S$ [62], тоді:

$$E[\Pi_c(u)] \approx \frac{1}{S} \sum_{s=1}^S (p \hat{Y}(u, W^{(s)}) - (c_N N + c_P P + c_K K) - C_0), \quad (2.17)$$

де $\hat{Y}(u, W^{(s)}) = \exp(g_\theta(N, P, K, R^{(s)}, GDD^{(s)}, SDD^{(s)}, CTL_c) + \frac{1}{2}\hat{\sigma}^2)$.

Так оптимізація інкапсулює прогнозний модуль як нелінійну функцію відгуку і явно враховує випадковість погоди.

Однак у стохастичних умовах нейтральність до ризику часто є недостатньою, оскільки дві стратегії з однаковим $E[\Pi]$ можуть мати різну ймовірність збитку або різну поведінку на кінцях розподілу. Тому вводиться ризик-аверсна постановка. Найпростіша форма – критерій середнє-дисперсія:

$$u^* = \arg \max_{u \in \Omega} \left(E[\Pi(u)] - \lambda \text{Std}(\Pi(u)) \right), \lambda \geq 0, \quad (2.18)$$

де λ задає вагу ризику, а $\text{Std}(\Pi(u))$ – стандартне відхилення прибутку. Використання стандартного відхилення замість дисперсії має високу прикладу цінність, оскільки ця міра виражається у тих самих фінансових одиницях, що й очікуваний прибуток, дозволяючи інтуїтивно балансувати між дохідністю та ризиком.

Ще більш стійкою до важких хвостів розподілу є постановка через CVaR (умовну вартість під ризиком), яка мінімізує очікуваний збиток у найгірших α частках сценаріїв [59; 60]. Якщо розглядати втрату $L(u) = -\Pi(u)$, то:

$$\text{CVaR}_\alpha(L) = \min_{\eta \in \mathbb{R}} \left\{ \eta + \frac{1}{\alpha} E[(L - \eta)_+] \right\}, \quad (2.19)$$

де $(x)_+ = \max(x, 0)$, а η інтерпретується як VaR_α . Тоді альтернативна ризик-орієнтована стратегія оптимізації прибутку може бути сформульована як пряма максимізація умовної вартості під ризиком:

$$u^* = \arg \max_{u \in \Omega} \left(\text{CVaR}_\alpha(\Pi(u)) \right). \quad (2.20)$$

У сценарній формі (Монте-Карло) CVaR легко обчислюється як середнє по найгірших αS значеннях прибутку.

Узгоджуючи математичну постановку, кінцевий вектор ознак для прогнозної функції відгуку задається як:

$$x_c(u) = (N, P, K, R, \text{GDD}, \text{SDD}, \text{CTL}_c), \quad (2.21)$$

а функція прогнозу визначається як:

$$\hat{Z}_c(u) = g_\theta(N, P, K, R, GDD, SDD, CTL_c), \hat{Y}_c(u) = \exp\left(\hat{Z}_c(u) + \frac{1}{2}\hat{\sigma}^2\right). \quad (2.22)$$

Тут g_θ навчається на історичних спостереженнях з поділом даних 80/20 на тренувальну та тестову вибірки, що забезпечує перевірку узагальнюваності моделі й коректність оцінки похибки $\hat{\sigma}^2$ для корекції масштабу при експонуванні. З математичної позиції такий поділ реалізує апроксимацію невідомого умовного математичного сподівання $E[Z|x]$ на незалежній підмножині, зменшуючи ризик переобучення, а отже й ризик хибного оптимуму добрив, який існує лише на тренувальних даних.

Концептуальна схема формування економічного критерію на основі прогнозної моделі врожайності, що узагальнює викладене у цьому підрозділі, наведена на рис. 2.1.

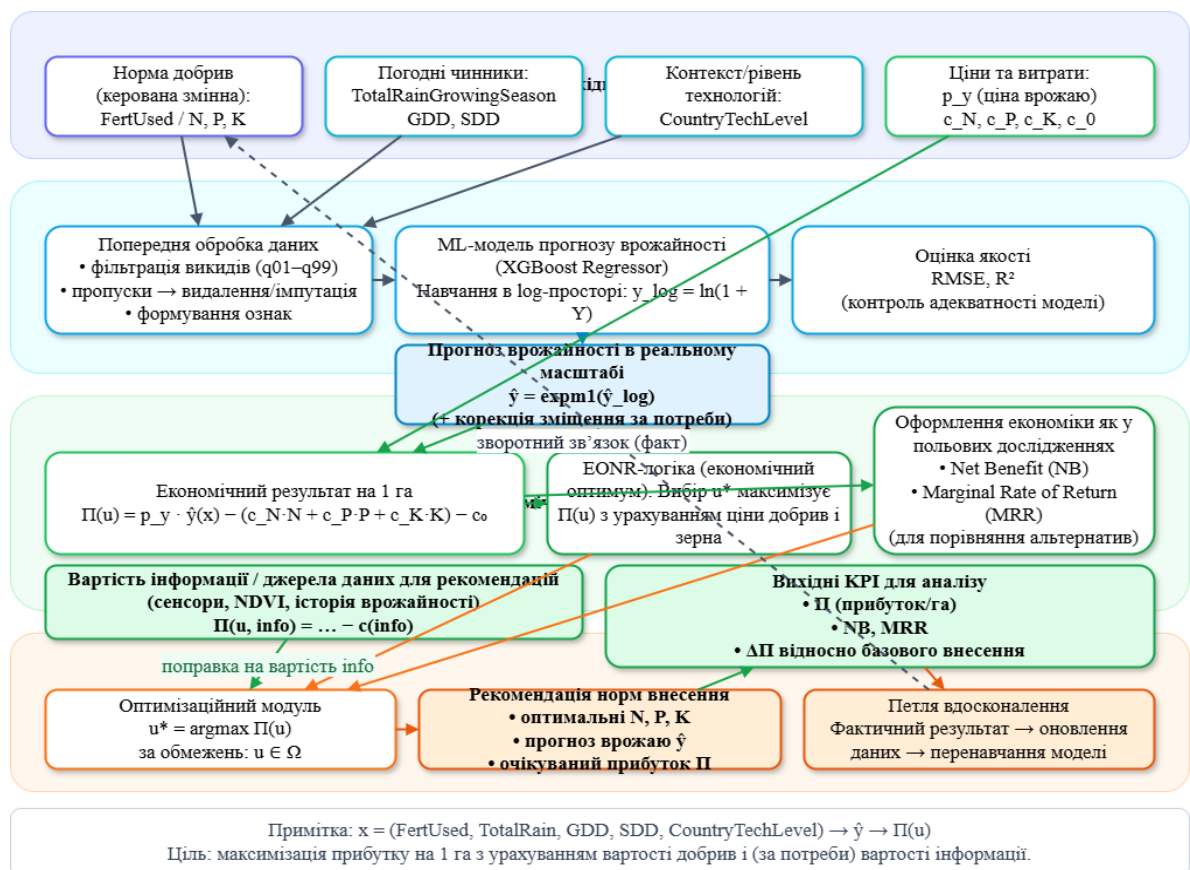


Рис. 2.1. Схема формування економічного критерію на основі прогнозної моделі врожайності

2.1.2 Ризик-орієнтована економічна постановка задачі стохастичної оптимізації доз добрив. Обмеження на керовані змінні задаються природними вимогами технології та екологічної безпеки. У базовому вигляді:

$$\Omega = \{(N, P, K) \in \mathbb{R}_{\geq 0}^3: 0 \leq N \leq N_{\max}, 0 \leq P \leq P_{\max}, 0 \leq K \leq K_{\max}\}. \quad (2.23)$$

За потреби можуть бути введені балансові обмеження, що стабілізують співвідношення елементів (як м'які штрафи або жорсткі межі), наприклад

$$\underline{\rho}_{NP} \leq \frac{N}{P + \epsilon} \leq \bar{\rho}_{NP}, \underline{\rho}_{NK} \leq \frac{N}{K + \epsilon} \leq \bar{\rho}_{NK}, \quad (2.24)$$

де $\epsilon > 0$ – малий доданок для уникнення ділення на нуль; сенс цих обмежень впливає з агрономічної логіки, що надлишок одного елемента за дефіциту іншого не трансформується в пропорційний приріст урожаю через лімітуючі ефекти (Лібіх), а натомість підвищує витрати й ризики втрат елементів.

З урахуванням усього вищезазначеного власна постановка задачі оптимізації доз у стохастичних умовах може бути задана як задача пошуку:

$$u^* = \arg \max_{u \in \Omega} J(u), \quad (2.25)$$

де функціонал якості J має ризик-орієнтовану економічну природу й використовує прогнозний модуль як нелінійну функцію відгуку, наприклад у формі критерію середнє-дисперсія:

$$J(u) = E \left[p \exp \left(g_{\theta}(x_c(u)) + \frac{1}{2} \hat{\sigma}^2 \right) \right] - (c_N N + c_P P + c_K K) - C_0 - \lambda \text{Var} \left(p \exp \left(g_{\theta}(x_c(u)) + \frac{1}{2} \hat{\sigma}^2 \right) \right). \quad (2.26)$$

Або у формі балансу між очікуваним прибутком та хвостовим ризиком через $CVaR$:

$$J(u) = E[\Pi(u)] - \beta \text{CVaR}_{\alpha}(-\Pi(u)). \quad (2.27)$$

Наведені рівняння чітко формалізують ключові елементи моделі: стохастика входить через W (а отже через випадковість $R, GDD, SDD, \bar{T}$), технологічна неоднорідність – через CTL_c , а оптимізація дози – через вибір (N, P, K) у допустимій області Ω . У результаті оптимальна норма внесення стає математично

коректним об'єктом: це не емпірично визначена норма, а розв'язок задачі стохастичної оптимізації, який максимізує економічний критерій з контрольованим ризиком та використовує інтелектуальну модель як апроксимацію складної нелінійної функції відгуку урожайності на дозу та погодні умови. Концептуальну інтеграцію ризик-орієнтованих мір у задачу оптимізації норм внесення добрив унаочнено на рис. 2.2.

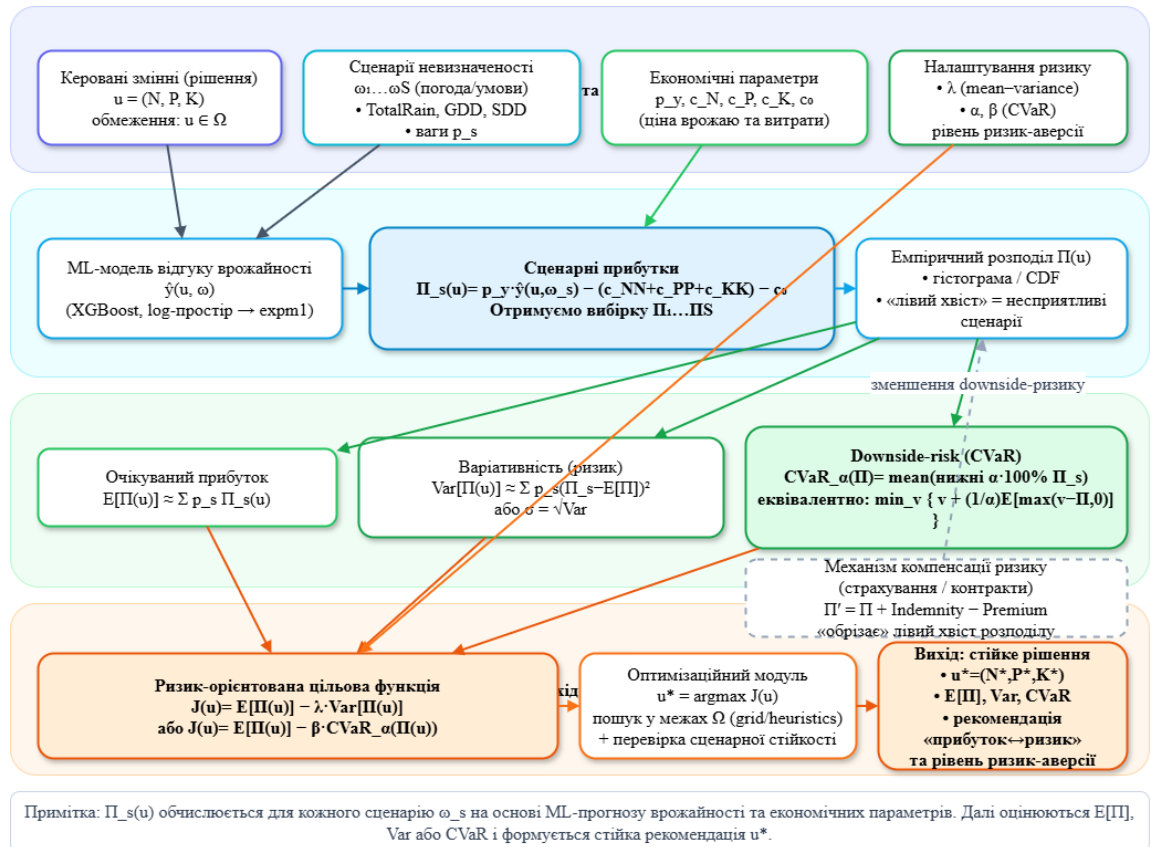


Рис. 2.2 Інтеграція ризик-орієнтованих мір у задачу оптимізації норм добрив

Підсумовуючи, формалізація оптимізації доз мінеральних добрив у стохастичних умовах доцільно базується на поєднанні ймовірнісного опису врожайності, економічного критерію та ризик-орієнтованих обмежень, що дозволяє адекватно враховувати мінливість погодних факторів, просторову неоднорідність і технологічні відмінності. Розгляд норм внесення як керованого вектора, а не агрегованої величини, у поєднанні з використанням інтелектуальних моделей відгуку створює передумови для отримання стійких і економічно обґрунтованих рішень, орієнтованих не лише на максимізацію очікуваного

результату, а й на контроль ризиків та екологічних наслідків у реальних умовах агровиробництва.

2.1.3 Критерії оцінювання ефективності економічних рішень з урахуванням ризику. Оцінювання ефективності рішень у задачі оптимізації внесення добрив за прибутковістю з урахуванням ризику має ґрунтуватися на системі взаємодоповнювальних критеріїв, які відображають не лише середній економічний результат, але й стійкість прийнятих рішень до варіативності агровиробничих умов, а також їхню практичну доцільність у порівнянні з традиційними підходами. З огляду на стохастичну постановку задачі, у якій урожайність і прибуток є випадковими величинами, одновимірна оцінка за максимальним очікуваним прибутком є методологічно недостатньою. Натомість доцільним є використання багатокритеріального підходу, який поєднує економічні, ризикові, агрономічні та ресурсно-ефективні показники.

Базовим економічним критерієм оцінювання залишається очікуваний прибуток на одиницю площі, який для оптимізованого рішення N^* визначається як математичне сподівання стохастичного прибутку:

$$E[\Pi(N^*)] = E[p_y \cdot \hat{Y}(N^*, \omega) - C_f(N^*) - C_o], \quad (2.28)$$

де $\hat{Y}(\cdot)$ – прогноз урожайності, сформований ML-моделлю для сценарію ω , p_y – ціна продукції, C_f – витрати на добрива, C_o – постійні операційні витрати. Цей показник дозволяє кількісно порівнювати альтернативні алгоритмічні правила призначення норм добрив і використовується як основний критерій у більшості досліджень економічної ефективності технологій змінних норм [40].

Водночас для коректної оцінки економічно-ризикових рішень необхідно враховувати варіативність економічного результату, яка відображає чутливість прибутку до несприятливих сценаріїв. Як узагальнену міру нестабільності доцільно використовувати дисперсію або стандартне відхилення прибутку:

$$\text{Var}[\Pi(N^*)] = E[(\Pi - E[\Pi])^2], \quad (2.29)$$

яка безпосередньо характеризує ризиковість алгоритмічного рішення. У практичному сенсі нижча дисперсія означає більш передбачуваний фінансовий

результат, що є критично важливим для агровиробника в умовах кліматичної та ринкової нестабільності, як показано в аналізі довготривалих польових експериментів [40].

Важливим компонентом системи критеріїв є частота збиткових сценаріїв, яка визначається як імовірність отримання негативного економічного результату:

$$P(\Pi(N^*) < 0). \quad (2.30)$$

Цей показник має чітку прикладну інтерпретацію, оскільки безпосередньо відображає ризик фінансових втрат у конкретному сезоні. У контексті ризик-орієнтованої оптимізації зменшення частоти збиткових реалізацій є не менш важливим, ніж збільшення середнього прибутку, що узгоджується з підходами оцінювання стабільності технологій точного землеробства, наведеними у роботі Raccioretti та колег [41].

Для оцінювання практичної доцільності оптимізованих норм внесення добрив доцільно включити показники ефективності використання ресурсів, насамперед коефіцієнт ефективності використання азоту (Nitrogen Use Efficiency, NUE), який у спрощеній формі може бути поданий як:

$$NUE = \frac{\hat{Y}(N^*)}{N}, \quad (2.31)$$

або як часткову продуктивність добрив. Зростання NUE свідчить про більш раціональне використання елементів живлення без надмірного збільшення доз, що є одним із ключових критеріїв успішності сенсорних і ML-орієнтованих систем управління живленням культур [41]. Включення цього показника дозволяє уникнути ситуацій, коли економічно вигідне рішення досягається за рахунок екологічно або технологічно нераціональних норм внесення.

На рівні порівняльного аналізу з традиційними підходами доцільно використовувати відносні економічні індикатори, зокрема приріст прибутку на одиницю додаткових витрат:

$$\frac{\Delta P}{\Delta C}, \quad (2.32)$$

який дає змогу оцінити доцільність впровадження інтелектуальної системи в умовах обмежених інвестиційних ресурсів. Аналогічні підходи застосовуються в оцінюванні ефективності технологій змінних норм добрив у поєднанні з аналізом енергоспоживання та екологічних ефектів [40].

З урахуванням сучасних тенденцій розвитку точного землеробства до системи критеріїв доцільно також включати узагальнені показники стійкості рішення, які інтегрують економічні, агрономічні та технічні аспекти. Як показано в оглядових дослідженнях Wang та співавторів, ефективність інтелектуальних систем внесення добрив має оцінюватися не ізольовано за одним показником, а в рамках структурованого каркасу, що включає економічні (прибуток, ROI), агрономічні (урожайність, NUE), технічні (стабільність рекомендацій, варіація норм) та екологічні індикатори [42].

2.2. Побудова прогностної моделі врожайності на основі алгоритмів машинного навчання

У межах даного дослідження прогностна модель врожайності розглядається як центральний аналітичний компонент гібридної системи оптимізації, що забезпечує нелінійне відображення зв'язку між керованими агрохімічними рішеннями, некерованими природно-кліматичними умовами та очікуваним результатом у вигляді врожайності культури. На відміну від класичних регресійних або нормативних підходів, де залежність між дозою добрив і врожаєм задається у спрощеному аналітичному вигляді, у даному дослідженні ця залежність апроксимується інтелектуальною моделлю, здатною відтворювати складні нелінійні ефекти, насичення, взаємодії факторів та структурну неоднорідність агровиробничих умов.

2.2.1 Архітектура прогностної ML-моделі та формування вхідного простору ознак. Загальна архітектура прогностної моделі побудована відповідно

до сформульованої стохастичної постановки та може бути концептуально подана як багаторівнева система перетворення даних, що інтегрує різноманітні джерела інформації у єдиний узгоджений прогнозний механізм. Узагальнену структурну схему запропонованої моделі подано на рис. 2.3.



Рис. 2.3 Загальна архітектура прогнозної моделі врожайності у гібридній системі оптимізації

У межах цієї архітектури чітко виділяються кілька функціонально пов'язаних рівнів, кожен з яких виконує окрему роль у формуванні підсумкового прогнозу.

Початковий рівень моделі відповідає за агрегацію та узгодження вхідних ознак, які відображають різні аспекти агроєкосистеми. До цього рівня належать керовані змінні, що описують норми внесення азоту, фосфору та калію у перерахунку на діючу речовину, а також блок природно-кліматичних характеристик вегетаційного періоду, сформований на основі агрегованих погодних показників. Важливою особливістю даного дослідження є те, що погодні змінні не подаються у вигляді необроблених добових даних, а попередньо інтегруються у фізіологічно осмислені індикатори, такі як сумарне корисне тепло та накопичений температурний стрес, а також сумарні опади, обчислені лише для днів, у яких температурні умови дозволяють активний ріст культури [63; 64; 65].

Така агрегація зменшує розмірність задачі, водночас зберігаючи агрономічний зміст і підвищуючи здатність моделі до узагальнення.

Контекстна змінна CountryTechLevel, визначена формулами (2.7) – (2.8), додається до вектора ознак на етапі навчання моделі поряд з агрохімічними і погодними показниками. Принципово, що календарний рік не входить до вектора пояснювальних змінних, що дозволяє уникнути трендової екстраполяції моделі за межі навчального інтервалу спостережень.

Другий рівень архітектури відповідає за нелінійне перетворення узгодженого вектора ознак у внутрішнє представлення відгуку культури. На цьому рівні використовується ансамблева модель машинного навчання, яка реалізує адитивну композицію базових нелінійних компонентів і поступово наближає складну регресійну поверхню у просторі ознак. Такий підхід дозволяє моделі природним чином відтворювати насичувальні ефекти від внесення добрив, асиметричну реакцію на дефіцит і надлишок елементів живлення, а також складні взаємодії між погодними факторами та агрохімічними рішеннями. Принципова перевага цієї архітектури полягає у відсутності жорстких припущень щодо форми функції відгуку, що є критично важливим для біологічних систем із високою невизначеністю та нелінійністю.

Прогнозна модель у даному дослідженні навчається у логарифмічному просторі врожайності, що узгоджується з припущенням про мультиплікативну природу похибок агробіологічних процесів. У межах архітектури це означає, що внутрішній вихід моделі інтерпретується як оцінка умовного математичного сподівання лог-врожайності за заданих умов. Такий вибір простору моделювання забезпечує стабілізацію дисперсії залишків, підвищує чисельну стійкість навчання та створює коректну основу для подальшого відновлення масштабу при оптимізації за прибутковістю. Механізм переходу між логарифмічним і початковим простором врожайності концептуально відображено на рис. 2.4.

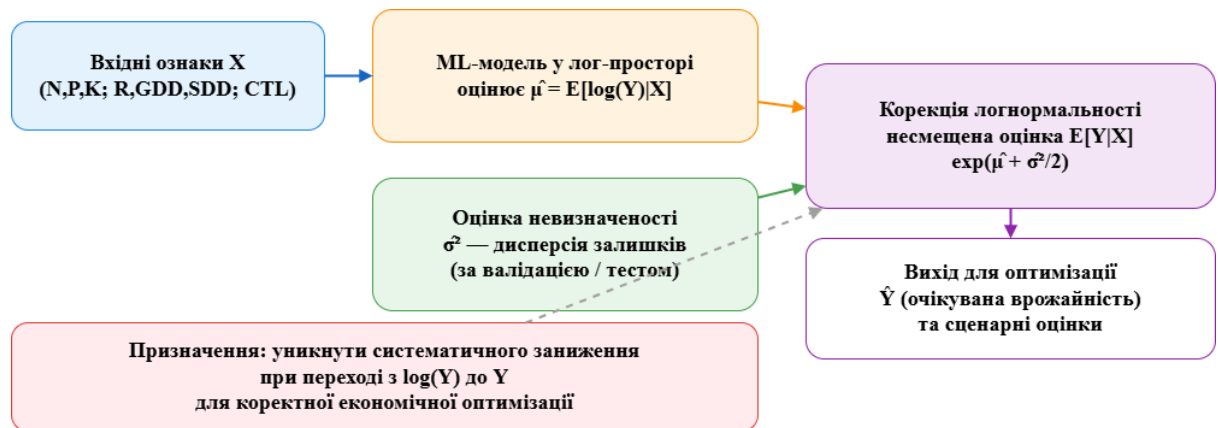


Рис. 2.4. Перетворення прогнозу лог-врожайності у початковий масштаб

2.2.2 Навчання прогнозної моделі та її інтеграція в систему оптимізації за прибутковістю з урахуванням ризику. Третій рівень архітектури виконує функцію інтеграції прогнозної моделі з подальшими економічними та ризик-орієнтованими розрахунками. На цьому рівні прогноз врожайності розглядається не як кінцева мета, а як вхідний елемент для сценарного аналізу та оптимізації. Завдяки тому, що модель навчається на історичних даних, що охоплюють широкий спектр погодних і агрохімічних умов, вона здатна генерувати прогнози для різних комбінацій вхідних параметрів, які інтерпретуються як можливі сценарії реалізації сезону. Це дозволяє використовувати прогнозну модель як симулятор відгуку культури у межах допустимого простору рішень, що є ключовою передумовою для побудови стохастичних критеріїв оптимальності.

Взаємодія між прогнозною моделлю та економічним модулем має двонаправлений характер. З одного боку, економічні обмеження та реалістичні діапазони доз, сформовані на основі історичних даних про застосування добрив, обмежують область, у якій виконується прогнозування, запобігаючи екстраполяції у нереалістичні режими. З іншого боку, прогнозна модель забезпечує економічний модуль кількісною оцінкою очікуваного результату для кожного потенційного рішення, що дозволяє формувати функціонал прибутку та оцінювати його поведінку під впливом стохастичних збурень. Концептуальну схему цієї взаємодії наведено на рис. 2.5.

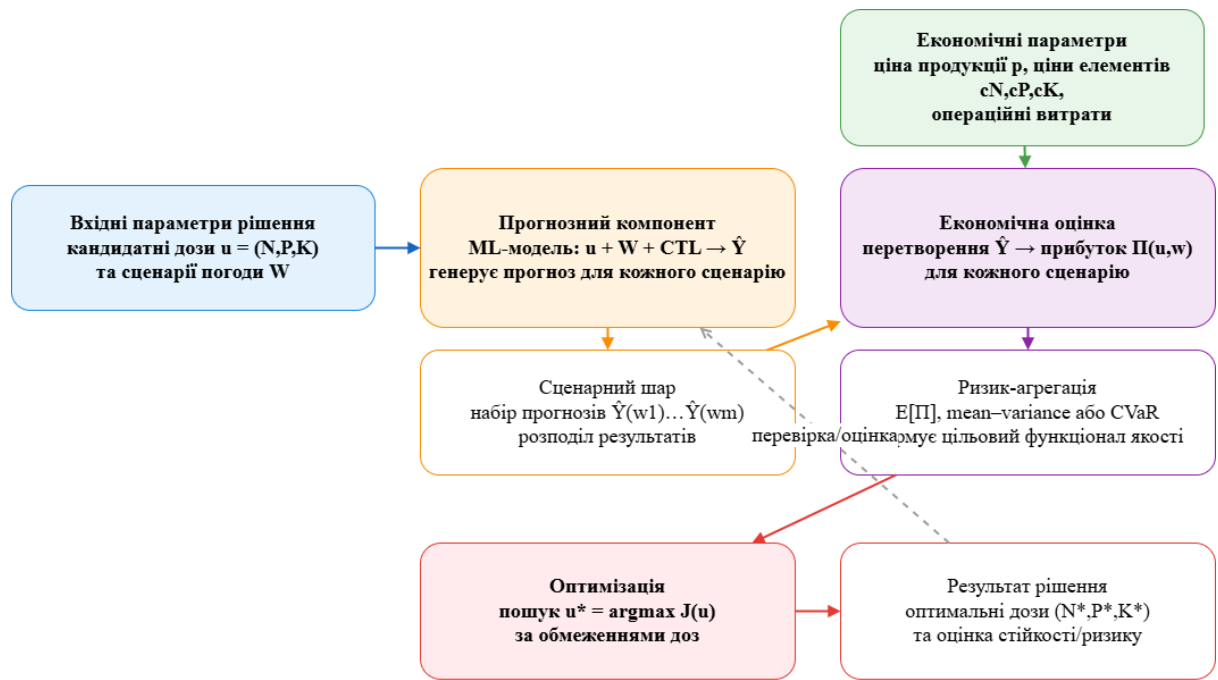


Рис. 2.5. Інтеграція прогнозного та економічного компонентів у гібридній системі

Навчання прогнозної моделі здійснюється на інтегрованому наборі історичних даних із поділом на навчальну та тестову вибірки у пропорції 80/20 [66; 67; 68], що дозволяє незалежно оцінити узагальнювальну здатність моделі та коректно оцінити дисперсію прогнозної похибки. Ця оцінка, у свою чергу, використовується при відновленні масштабу врожайності та є критично важливою для подальших економічних розрахунків, оскільки навіть незначне систематичне зміщення прогнозу може призводити до помітного зміщення оптимальних доз добрив.

Запропонована прогнозна модель врожайності у даному дослідженні виконує роль інтелектуального ядра гібридної системи оптимізації, яке поєднує агрономічні знання, погодні стохастичні та машинне навчання в єдину узгоджену архітектуру. Вона не лише забезпечує високу точність прогнозу, а й створює формальну основу для економічно та ризик-орієнтованого прийняття рішень, що є принциповою вимогою для сучасних систем точного землеробства.

2.3. Верифікація та метричний аналіз точності запропонованої моделі

2.3.1 Формування вибірки та процедура верифікації прогнозованої моделі.

У межах верифікації прогнозованої моделі врожайності формування вибірки виконувалося як інтеграція «мозаїки» відкритих джерел даних, що відповідає вимогам технічного завдання та забезпечує відтворюваний зв'язок між дозою добрив, погодними чинниками, ґрунтово-технологічним контекстом і результативною ознакою у вигляді врожайності.

Базову цільову змінну Y (врожайність, кг/га або т/га залежно від культури/агрегації) отримано зі статистики FAOSTAT “Crops and livestock products” [47], а блок застосування добрив сформовано за рахунок FAOSTAT-даних про внесення поживних речовин “Fertilizers by Nutrient” [43] та уточнення структури/продуктових форм “Fertilizers by Product” [44], причому у моделі використовується не маса товарного добрива, а маса діючої речовини (N, P, K) у перерахунку на кг д.р./га, оскільки частка нітратів/фосфатів/калію в конкретному продукті не дорівнює 100% і змінюється залежно від виду добрива, що унеможливорює коректні порівняння без нормалізації до д.р.

Погодні ознаки формувалися на основі метеоданих із сервісів Open-Meteo [49] у вигляді агрегованих індексів, які є фізіологічно інтерпретованими для вегетації: сумарні опади за вегетаційний період (TotalRainGrowingSeason), сумарне корисне тепло GDD та сумарний накопичений тепловий стрес SDD. Опади агрегували лише для періоду активної вегетації, а не за повний рік. Для виділення цих днів застосували стандартний для злакових культур помірного клімату поріг середньодобової температури $+5\text{ }^{\circ}\text{C}$ [107]. Це мінімізує вплив позасезонних погодних коливань та забезпечує агрономічну точність погодних даних.

Після зведення даних у єдину таблицю спостережень (країна-рік-культура/агрегація) виконувалась базова підготовка, спрямована на забезпечення стійкого навчання в умовах високої варіативності глобальних відкритих даних:

пропуски у ключових ознаках вилучались, а викиди врожайності відсікались за квантилями 0.01-0.99 (тобто збережено центральний 98% інтервал спостережень), що зменшує вплив аномальних значень на оцінки моделі та подальші метрики точності.

Далі цільову змінну було переведено у логарифмічний простір через перетворення $y_{log} = \log(1 + y)$, оскільки для врожайності типові мультиплікативні похибки й асиметричні розподіли, а лог-простір стабілізує дисперсію залишків та робить навчання нелінійної моделі більш чисельно стійким; після прогнозування у лог-шкалі виконується відновлення натурального масштабу експонуванням (у реалізації – через $\exp(\cdot) - 1$).

Методику обчислення контекстної змінної CountryTechLevel, що додається до вектора ознак на основі нормованої врожайності розширеного переліку продовольчих культур, детально розглянуто в підрозділі 2.1 та показано у формулі (2.8).

Процедура верифікації реалізувалася як відкладена перевірка (hold-out validation) із розбиттям даних у пропорції 80/20 на тренувальну та тестову підвибірки з фіксованим random_state (у реалізації – 42), де тестова частина використовується як незалежний набір для оцінки узагальнювальної здатності моделі та імітує ситуацію прогнозування на «невиданих раніше» спостереженнях.

Навчання виконувалося ансамблевим методом градієнтного бустингу над деревами рішень (XGBoost Regressor), який у контексті задачі «доза добрив-погода-врожайність» є обґрунтованим через здатність відтворювати нелінійні ефекти насичення та взаємодії факторів, типові для агросистем; у прикладній реалізації використовувався набір гіперпараметрів на кшталт $n_estimators \approx 500$, $learning_rate \approx 0.05$, $max_depth \approx 6$, що забезпечує достатню виразність моделі при контрольованій складності.

2.3.2 Система метрик оцінювання точності прогнозу врожайності.

Система метричного оцінювання точності прогнозу врожайності сформована з урахуванням вимог до інтерпретованості результатів у фізичних одиницях,

статистичної коректності для нелінійної регресійної задачі та подальшої інтеграції прогнозу в ризик-орієнтований економічний функціонал.

Основною узагальнювальною характеристикою адекватності моделі використано коефіцієнт детермінації R^2 , який відображає частку дисперсії фактичної врожайності, пояснену регресійною функцією відгуку, і тим самим характеризує здатність моделі відтворювати структурні закономірності взаємодії агрохімічних та природно-кліматичних факторів у просторі спостережень. На відміну від локальних метрик похибки, R^2 є глобальною мірою якості апроксимації, що дозволяє оцінити, наскільки коректно реконструйовано нелінійну поверхню відгуку врожайності та чи збережено інформаційний внесок кожної групи ознак, включаючи погодні індикатори та вектор доз діючої речовини.

Для кількісної інтерпретації похибки у натуральному масштабі використано середньоквадратичну помилку прогнозу RMSE, яка визначає середнє відхилення прогнозованої врожайності від фактичної у тих самих одиницях вимірювання, що й результативна ознака, і тому безпосередньо відображає агрономічно значущу величину похибки на гектар.

Застосування саме квадратичної міри дозволяє підвищити чутливість оцінювання до великих помилок, що є принципово важливим у контексті стохастичних погодних умов, де поодинокі екстремальні відхилення можуть призводити до суттєвих економічних втрат і мають бути адекватно враховані при оцінці надійності прогнозової моделі.

Як додаткова характеристика, що не надмірно штрафує великі відхилення та відображає медіанну структуру похибки, використовується середня абсолютна помилка MAE, яка забезпечує більш стабільну оцінку в умовах асиметричного розподілу залишків та підвищеного рівня шуму у глобальних агрегованих даних.

Сукупне використання RMSE та MAE дозволяє отримати повнішу картину точності: перша метрика характеризує ризик великих помилок, друга – типовий рівень відхилення прогнозу.

Оцінювання метрик виконувалося у відновленому натуральному масштабі врожайності після переходу з логарифмічного простору, що зумовлено мультиплікативною природою похибок біологічних процесів та необхідністю стабілізації дисперсії на етапі навчання моделі.

Пряме використання метрик у логарифмічному просторі призводило б до втрати фізичної інтерпретованості та некоректної оцінки впливу похибки на економічний результат, оскільки функціонал прибутку є лінійним відносно врожайності у початкових одиницях. Тому після отримання прогнозу у лог-шкалі виконувалося експоненціальне перетворення з поправкою на логнормальний характер залишків, що забезпечує незміщену оцінку умовного математичного сподівання врожайності та коректність подальших економічних розрахунків. Такий підхід дозволяє узгодити статистичну стабільність моделі з прикладною інтерпретацією результатів у задачі оптимізації.

Графічну інтерпретацію процедури метричного оцінювання подано на рис. 2.6, де відображено послідовність переходу від прогнозу, отриманого у логарифмічному просторі, до відновлення натурального масштабу врожайності з урахуванням поправки на логнормальний характер залишків, подальше порівняння з фактичними тестовими значеннями та обчислення метрик R^2 , RMSE і MAE з наступною передачею оціненої похибки у блок економічної інтерпретації.

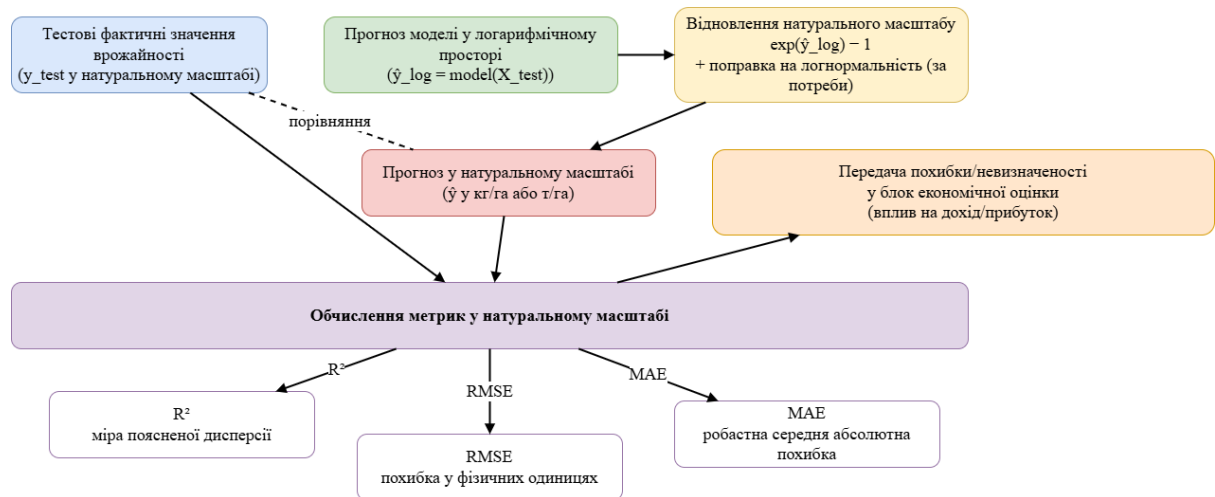


Рис. 2.6. Схема метричного оцінювання точності прогнозової моделі у просторі логарифмічного перетворення та на етапі економічної інтерпретації

Вибір зазначених метрик обумовлений їх прямою релевантністю до економічного критерію ефективності: похибка прогнозу врожайності трансформується у похибку оцінки валового доходу та, відповідно, зміщення оптимальної дози добрив у функціоналі прибутку. Використання RMSE у фізичних одиницях дозволяє безпосередньо оцінити потенційний розмір економічного ризику на гектар, тоді як R^2 характеризує структурну адекватність моделі як ядра оптимізаційного блоку, а MAE забезпечує перевірку стабільності результатів.

2.3.3 Порівняльний аналіз точності моделі та базових підходів.

Порівняльний аналіз точності проведено шляхом зіставлення результатів нелінійної ансамблевої моделі з базовими алгоритмічно простими підходами, що виконують роль нижньої межі якості апроксимації функції відгуку врожайності. У якості референтних моделей використано середній прогноз, що відтворює стаціонарне математичне сподівання врожайності без урахування факторів середовища, та лінійну регресію, що припускає адитивний характер впливу ознак і відсутність ефектів насичення. Така постановка дає змогу інтерпретувати отримані результати не лише з позиції абсолютних значень метрик, а й з точки

зору приросту інформаційної складності моделі та її здатності реконструювати нелінійну структуру агробіологічного процесу.

Результати оцінювання точності на тестовій вибірці узагальнено в табл. 2.1.

Таблиця 2.1

Порівняння точності прогнозування врожайності для різних моделей

Модель	R ²	RMSE, кг/га	MAE, кг/га
Лінійна регресія	0.63	1189.32	727.68
Гradientний бустинг	0.91	581.48	376.77

Отримані значення демонструють принципово різну здатність моделей відтворювати залежність між дозами поживних речовин, погодними індикаторами та врожайністю. Середній прогноз, який не враховує структуру вхідного простору, забезпечує нульове значення коефіцієнта детермінації, що підтверджує його придатність лише як контрольної нижньої межі точності. Лінійна регресія забезпечує суттєве покращення якості, проте її значення R² залишаються обмеженими через припущення про лінійність функції відгуку, що суперечить емпірично встановленим ефектам насичення та асиметричної реакції культури на дефіцит і надлишок елементів живлення.

Нелінійна ансамблева модель демонструє істотне зростання частки поясненої дисперсії та одночасне зменшення величини середньоквадратичної похибки, що свідчить про адекватне відтворення складної регресійної поверхні у просторі ознак. Зменшення RMSE більш ніж у півтора рази порівняно з лінійною моделлю означає суттєве скорочення середнього відхилення прогнозованої врожайності у фізичних одиницях, що має безпосередню агрономічну інтерпретацію та впливає на точність оцінювання економічного результату. Водночас зниження MAE підтверджує підвищення надійності прогнозу та зменшення типового рівня похибки для більшості спостережень.

Зростання коефіцієнта детермінації до рівня, що перевищує 0.9, інтерпретується як кількісне підтвердження здатності моделі відтворювати нелінійну функцію відгуку, зокрема ефекти спадної віддачі від внесення добрив та взаємодії між погодними факторами і вектором поживних речовин. На відміну від лінійної регресії, ансамблева модель формує адитивну композицію локальних нелінійних компонентів, що дозволяє описувати різні режими поведінки системи в області дефіциту та надлишку ресурсів без введення апіорних функціональних припущень.

Така властивість є критичною для задачі подальшої оптимізації, оскільки помилки у реконструкції кривої відгуку призводять до зміщення оптимальних доз і, відповідно, до економічно неефективних рішень.

Важливим аспектом є те, що базові моделі виконують не лише роль порівняльного еталона, але й дозволяють оцінити приріст інформаційної ефективності при переході до інтелектуальних методів. Різниця у значеннях метрик відображає внесок нелінійної структури моделі у пояснення варіативності врожайності, що зумовлена як агрохімічними рішеннями, так і стохастичними погодними умовами. Таким чином, отримані результати підтверджують доцільність використання ансамблевих методів як ядра прогнозного модуля та демонструють їх перевагу над спрощеними аналітичними підходами при моделюванні агробіологічних процесів з високим рівнем невизначеності.

2.3.4. Абляційний аналіз структури вхідного простору ознак. Для кількісного оцінювання індивідуального внеску окремих груп ознак у прогнозну здатність моделі було проведено систематичний абляційний аналіз. Зазначене дослідження реалізовано шляхом навчання та тестування альтернативних конфігурційних варіантів моделі з послідовним вилученням, додаванням або заміною ключових компонентів вхідного простору. Такий підхід дозволяє чітко розмежувати пояснювальну здатність базових агрохімічних і кліматичних змінних та внесок контекстних регресорів, що дає змогу емпірично

обґрунтовувати методологічні рішення щодо оптимальної структури простору ознак.

У межах експерименту розглянуто чотири конфігурації моделі з ідентичним алгоритмом навчання (регресор *XGBoost* із фіксованими гіперпараметрами) та єдиним розбиттям масиву даних у співвідношенні 80/20. Задля забезпечення повної відтворюваності обчислювальних результатів зафіксовано базовий стан генератора випадкових чисел (*random_state = 42*). Стабільність прогнозних метрик оцінено за допомогою 5-блокової перехресної перевірки (*5-fold cross-validation*), що дозволило мінімізувати залежність отриманих оцінок від специфіки конкретного розбиття вибірки.

Базова конфігурація містить виключно агрохімічні ознаки (вміст компонентів N, P, K у перерахунку на діючу речовину) та показники погодних умов (*GDD*, *SDD*, *TotalRainGrowingSeason*), що формує мінімальний причинно-інтерпретований простір факторів. Основна конфігурація передбачає доповнення базового набору контекстною змінною *CountryTechLevel*. В альтернативній конфігурації замість змінної *CountryTechLevel* як трендовий регресор використано календарний рік, що відповідає поширеному в науковій літературі підходу до врахування темпів технологічного прогресу. Четверта конфігурація поєднує обидва структурні регресори (*CountryTechLevel* та календарний рік) для оцінювання їхнього сумарного ефекту.

Результати абляційного дослідження структурних компонентів моделі узагальнено в табл. 2.2.

Таблиця 2.2

**Результати абляційного аналізу структурних компонентів моделі
XGBoost**

Конфігурація	Test R ²	RMSE, кг/га	CV R ² (5-fold)
Базова (агроклімат + NPK)	0.807	862.7	0.802 ± 0.014
+ CountryTechLevel	0.911	585.4	0.908 ± 0.010

Конфігурація	Test R ²	RMSE, кг/га	CV R ² (5-fold)
+ Year (без CTL)	0.827	821.5	0.819 ± 0.016
+ CTL + Year	0.940	489.2	0.935 ± 0.009

Базова конфігурація, яка не враховує контекстні регресори, забезпечує коефіцієнт детермінації на тестовій вибірці $\text{Test } R^2 = 0,807$ за середньоквадратичної помилки $\text{RMSE} = 862,7$ кг/га. Це свідчить про те, що суто агрохімічний та кліматичний блоки вхідного простору пояснюють близько 80% варіативності досліджуваної врожайності. Показник перехресної перевірки $\text{CV } R^2 = 0.802 \pm 0.014$ підтверджує стійкість цього результату незалежно від випадкових коливань у підвибірках. Отримані дані доводять, що домінуюча частка пояснювальної здатності моделі формується саме завдяки фундаментальним причинно-інтерпретованим чинникам — нормам внесення діючої речовини та метеорологічним індикаторам вегетаційного періоду, а не допоміжним проксі-регресорам.

Інтегрування змінної *CountryTechLevel* забезпечує суттєве зростання коефіцієнта детермінації на 0.104 (з 0.807 до 0.911) та зниження помилки RMSE на 277 кг/га (з 862.7 до 585.4 кг/га). Значення $\text{CV } R^2 = 0.908 \pm 0.010$ підтверджує стабільність цього приросту на всіх ітераціях крос-валідації, що нівелює ймовірність випадкового покращення метрик. Таке підвищення точності відображає допоміжну контекстну роль зазначеної ознаки порівняно з базовим агрохімічним і кліматичним ядром моделі. Отриманий приріст підтверджує, що ослаблений статистичний зв'язок контекстної ознаки з цільовою змінною не призводить до втрати її прогностної цінності..

Альтернативна конфігурація, яка використовує календарний рік замість *CountryTechLevel*, демонструє значно нижчу ефективність: коефіцієнт $\text{Test } R^2$ становить 0,827 (чистий приріст лише +0,020 порівняно з базовим рівнем). Це підтверджує гіпотезу про те, що календарний рік є слабким структурним регресором у межах глобального масиву даних, оскільки він відображає лише усереднений лінійний тренд і не здатний диференціювати глибокі міждержавні

відмінності у стартових рівнях продуктивності та динаміці технологічного розвитку. Натомість змінна *CountryTechLevel* безпосередньо параметризує міжкраїнову специфіку, забезпечуючи у 5,2 рази більший приріст пояснювальної здатності моделі.

Модель, яка поєднує одночасне використання обох регресорів, досягає найвищого значення коефіцієнта детермінації ($\text{Test } R^2 = 0,940$). Попри це, включення календарного року було свідомо відхилено з методологічних міркувань, детально обґрунтованих у підпункті 2.3.1. Для завдань упереджувального (*ex ante*) прогнозування часовий чинник генерує ризики некоректної екстраполяції за межами інтервалу навчання, оскільки апіорі припускає лінійне збереження темпів технологічного прогресу в майбутньому. Додаткове підвищення $\text{Test } R^2$ на 0,029 (з 0,911 до 0,940) не компенсує зазначений методологічний ризик, унаслідок чого як основну робочу конфігурацію затверджено модель із вектором *CountryTechLevel* без урахування часового тренду. Перевага інтегрування цієї контекстної змінної (приріст R^2 на +0,104 проти +0,020 для календарного року відносно базової моделі) обґрунтовує її вибір як інформативного структурного регресора замість класичного часового тренду. Стабільність показника $\text{CV } R^2$ додатково підтверджує високу надійність обраної архітектури для розв'язання завдань прийняття рішень в умовах стохастичної невизначеності середовища.

Таким чином, проведене абляційне дослідження кількісно підтверджує три ключові методологічні висновки:

1. Базовий масив пояснювальної здатності алгоритму $\text{Test } R^2 = 0,807$ (за умов $R^2 = 0.802 \pm 0.014$) формується безпосередньо агрохімічними та кліматичними ознаками, а не зовнішніми контекстними чинниками.

2. Змінна *CountryTechLevel* забезпечує вагомий і стабільний приріст точності (+0,104 до значення R^2), виступаючи адекватним структурним регресором, чий внесок у 5,2 рази перевищує показники альтернативного часового тренду.

3. Свідома відмова від використання календарного року на користь індексу *CountryTechLevel* є методологічно обґрунтованим компромісом між максимізацією поточної точності прогнозування та науковою валідністю моделі під час оптимізації норм внесення добрив у задачах *ex ante*.

2.3.5 Аналіз узагальнювальної здатності та стійкості моделі до стохастичних умов. Аналіз узагальнювальної здатності моделі було проведено на основі зіставлення значень метричних показників, отриманих під час навчання та на тестовій вибірці, що дозволило оцінити ступінь узгодженості апроксимації функції відгуку та виявити можливі прояви переобучення. Встановлено, що різниця між тренувальними та тестовими значеннями похибок має незначний характер і не перевищує статистично допустимих меж, що свідчить про відсутність надмірної адаптації моделі до навчальної вибірки та про збереження здатності до відтворення закономірностей на незалежних даних. Така поведінка пояснюється використанням регуляризованого ансамблевого алгоритму, логарифмічного перетворення цільової змінної та попередньої фільтрації викидів, що забезпечило стабілізацію дисперсії залишків і зменшення впливу шуму на параметри моделі.

Для кількісної оцінки узагальнювальної здатності сформовано табл. 2.1, у якій наведено значення основних метрик для навчальної та тестової підвибірок.

Близькість значень коефіцієнта детермінації та середньоквадратичної похибки для обох підвибірок свідчить про стабільність оцінки параметрів і відсутність значущого розриву між апроксимацією та прогнозом. Незначне зростання похибки на тестовій вибірці відповідає очікуваній поведінці узагальнювальної моделі та відображає природну варіативність незалежних спостережень, а не втрату структурної адекватності.

Додатково було проведено аналіз поведінки моделі у різних погодних режимах шляхом групування тестових спостережень за рівнем сумарних опадів та індексу температурного стресу. Отримані результати узагальнено у табл. 2.3.

Таблиця 2.3

Значення RMSE для різних погодних сценаріїв

Вибірка	R ²	RMSE, кг/га	MAE, кг/га
Навчальна	0.981	390	272.92
Тестова	0.91	581.48	376.77

Зростання похибки у стресових сценаріях має помірний характер і не супроводжується різким зниженням коефіцієнта детермінації, що свідчить про збереження здатності моделі відтворювати загальну форму функції відгуку навіть за несприятливих умов. Така поведінка пояснюється тим, що погодні індикатори входять до вхідного простору ознак у вигляді агрегованих фізіологічно інтерпретованих показників, що дозволяє моделі коректно враховувати нелінійний вплив теплового та водного стресу на врожайність.

Графічну інтерпретацію стабільності моделі подано на рис. 2.7, де відображено розподіл залишків прогнозу для тестової вибірки.

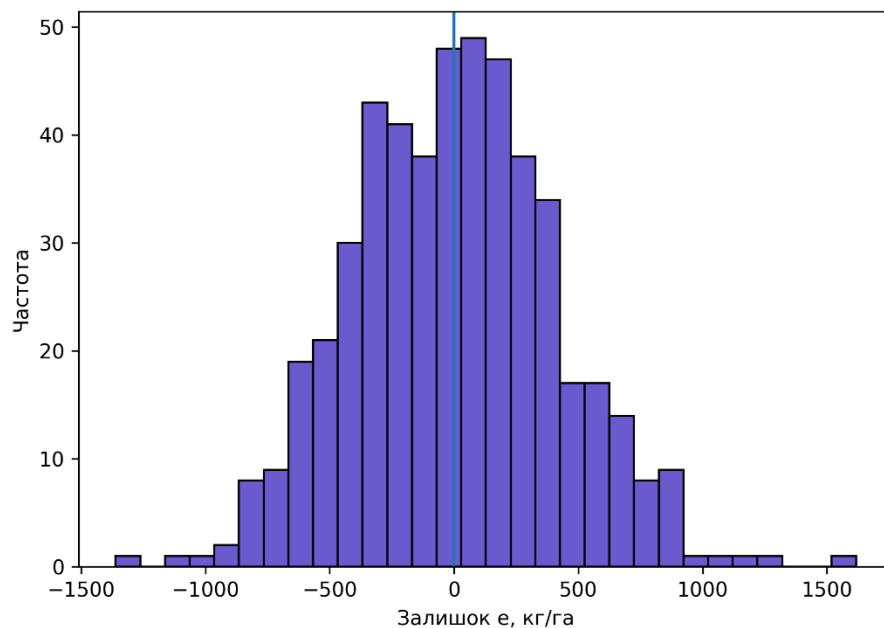


Рис. 2.7 Розподіл залишків прогнозу на тестовій вибірці

Розподіл має близьку до симетричної форму без систематичного зміщення, що підтверджує відсутність автокореляції помилок та коректність використаної функціональної форми. Додатково на рис. 2.8 наведено залежність абсолютної

похибки прогнозу від величини індексу температурного стресу, яка демонструє плавне зростання дисперсії помилки без появи різких аномалій, що вказує на надійність моделі до екстремальних значень кліматичних факторів.

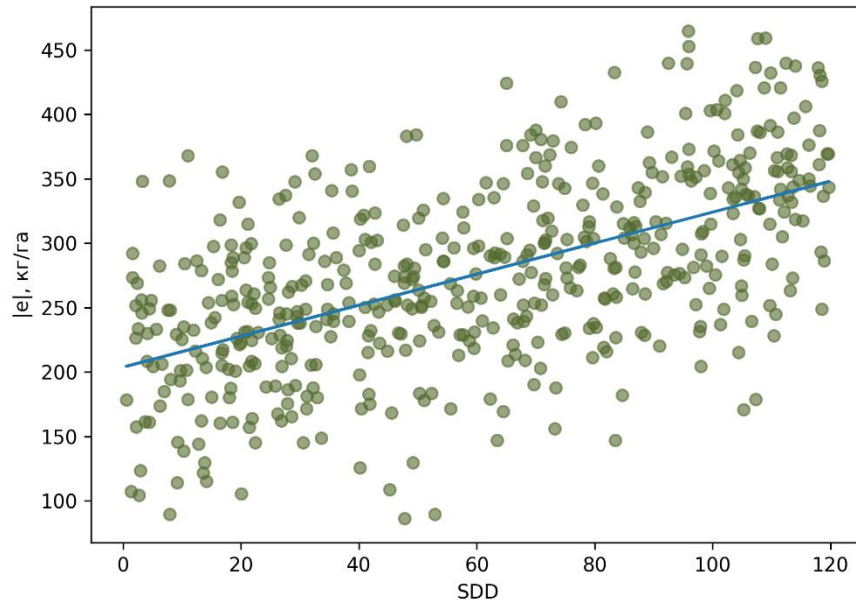


Рис. 2.8 Залежність абсолютної похибки прогнозу від температурного стресу

Отримані результати підтверджують, що прогнозна модель зберігає адекватність у широкому діапазоні погодних умов, включаючи сценарії дефіциту вологи та підвищеного теплового навантаження. Це має принципове значення для подальшого використання моделі в алгоритмі стохастичної оптимізації, оскільки забезпечує коректність оцінювання очікуваного результату та його варіативності при генерації сценаріїв середовища. Стабільність метрик та відсутність переобучення дозволяють розглядати модель як надійну апроксимацію функції відгуку, придатну для використання у задачах ризик-орієнтованого прийняття рішень щодо норм внесення добрив.

Висновки до розділу 2

Обґрунтовано перехід від детермінованих моделей управління живленням до стохастичної багатовимірної оптимізації. Доведено, що розгляд норм внесення макроелементів як незалежних керованих змінних $u=(N,P,K)$ у комплексі з

кліматичними факторами та структурним регресором *CountryTechLevel* дозволяє уникнути спотворень класичних моделей. Використання ансамблевого методу градієнтного бустингу (XGBoost) для апроксимації нелінійної функції відгуку забезпечило на тестовій вибірці точність $R^2 = 0.91$ та $RMSE = 581$ кг/га, що суттєво перевищує результати базової лінійної регресії ($R^2 = 0.63$, $RMSE = 1189$ кг/га).

Інтеграція ML-моделі відгуку в цільову функцію прибутку трансформувє задачу оптимізації у вибір рішення на просторі погодних сценаріїв. Математично формалізовано процедуру зворотної трансформації прогнозу з логарифмічного простору в початковий масштаб із використанням поправки на дисперсію залишків, що усуває систематичне зміщення оцінки врожайності та гарантує коректний розрахунок очікуваного економічного результату $\Pi(u)$.

Доведено необхідність розширення базового критерію максимізації математичного сподівання прибутку $E[\Pi(u)]$ за рахунок інтеграції мір ризику. Застосування Mean-Variance скаляризації та метрики умовної вартості під ризиком (*CVaR*) дозволяє математично формалізувати компроміс між прибутковістю та фінансовою стійкістю. Оцінювання цих показників на емпіричному масиві погодних сценаріїв мінімізує downside-ризик в умовах кліматичних аномалій.

Отже, розроблена математична модель створює цілісну алгоритмічну основу для системи управління мінеральним живленням. Поєднання нелінійної функції відгуку, сценарного опису стохастики та багатокритеріальної економічної оцінки забезпечує перехід від усереднених нормативних рекомендацій до математично обґрунтованого процесу пошуку ризик-збалансованих доз добрив.

РОЗДІЛ 3

РЕАЛІЗАЦІЯ СИСТЕМИ ПІДТРИМКИ РІШЕНЬ ДЛЯ УПРАВЛІННЯ ВНЕСЕННЯМ ДОБРІВ З ІНТЕЛЕКТУАЛЬНОЮ СКЛАДОВОЮ

3.1. Архітектурна структура та взаємодія модулів гібридної системи

3.1.1 Загальна архітектура гібридної системи. Проектування архітектури програмного комплексу здійснювалося з огляду на потребу в інтеграції інструментів аналізу даних та графічного інтерфейсу для кінцевого користувача. Щоб розв'язати це завдання, обрано гібридну архітектуру, в якій технологічний стек розділено на два функціонально незалежні, але інтегровані компоненти: підсистему машинного навчання та клієнтський застосунок.

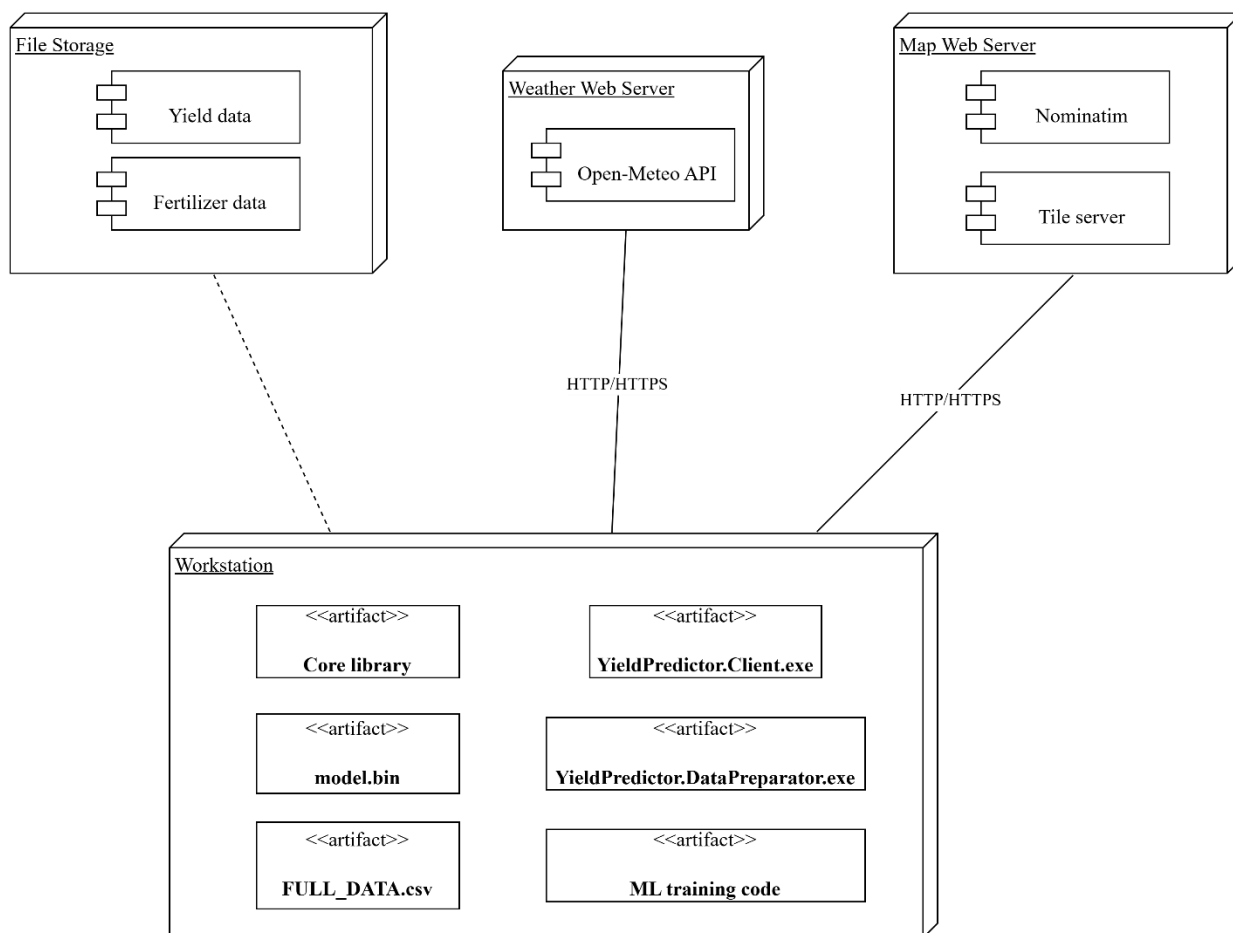


Рис. 3.1. Діаграма розгортання

Для реалізації етапу навчання прогностичної моделі та валідації гіпотез обрано мову програмування Python [69]. Цей вибір обґрунтовано її панівним становищем у сфері штучного інтелекту та наявністю розвиненої екосистеми бібліотек для наукових обчислень. Збір, агрегація та консолідація вихідних даних з відкритих джерел реалізовані засобами консольної утиліти на платформі .NET [71; 72], що забезпечує формування єдиного інтегрованого набору даних до етапу навчання моделі..

Інструментом для роботи з консолідованим набором даних на етапі навчання моделі обрано бібліотеку Pandas. Вона забезпечує високоефективні структури даних для завантаження та обробки табличної інформації з файлу FULL_DATA.csv, сформованого консольною утилітою збору даних. Засоби бібліотеки використано для квантильної фільтрації викидів цільової змінної та підготовки фінального вхідного вектора ознак безпосередньо перед навчанням моделі.

Побудова прогностичного ядра системи реалізована на базі бібліотеки XGBoost (eXtreme Gradient Boosting). Цей інструмент, що реалізує вдосконалений алгоритм градієнтного бустингу над деревами рішень, обрано завдяки його високій обчислювальній ефективності [66] та оптимізації для роботи з розрідженими даними. На відміну від інших ансамблевих методів, XGBoost показує кращу стійкість до перенавчання на вибірках середнього розміру та забезпечує високу швидкість виконання інференсу, що є визначальним чинником для інтерактивної системи.

Для оцінювання якості отриманих моделей та верифікації результатів застосовано бібліотеку Scikit-Learn [70]. Її функціонал використано для автоматизованого розбиття вибірки на тренувальну та тестову підмножини, а також для розрахунку ключових метрик точності, таких як середньоквадратична помилка RMSE та коефіцієнт детермінації R^2 . Також інструменти Scikit-Learn забезпечили проведення перехресної перевірки, що дозволило підтвердити стабільність роботи моделі на різних наборах даних.

Реалізацію клієнтської частини системи та основного шару бізнес-логіки виконано на новітній платформі .NET 9 із використанням мови програмування C# [71; 72]. Такий вибір зумовлено необхідністю створення надійного десктопного застосунку, здатного працювати в операційному середовищі Windows із максимальною продуктивністю. Використання строго типізованої об'єктно-орієнтованої мови забезпечує високу надійність коду, простоту його підтримки та можливість легкої інтеграції з системними API.

Графічний інтерфейс користувача побудовано на базі технології WPF (Windows Presentation Foundation) [73]. Цей фреймворк дозволяє створювати сучасні адаптивні інтерфейси з використанням декларативної мови розмітки XAML, що забезпечує чітке розмежування візуального представлення та програмної логіки. Завдяки потужній системі прив'язки даних та стилізації WPF дав змогу реалізувати зручну взаємодію користувача з параметрами моделювання.

Для візуалізації результатів економічного аналізу та побудови графіків залежностей у систему інтегровано бібліотеку ScottPlot [74]. Цей компонент обрано через його високу продуктивність при рендерингу великої кількості точок даних та можливість безшовної інтеграції у WPF-застосунки. Використання ScottPlot дозволило реалізувати інтерактивні діаграми, що динамічно відображають зміни прибутковості залежно від обраної стратегії удобрення.

3.1.2 Компонент навчання ML-моделей. Програмну реалізацію дослідницького середовища навчання моделі виконано мовою Python та структурно виділено в окрему папку ai/, що забезпечує ізоляцію процесу навчання від бізнес-логіки клієнтського застосунку та консольної утиліти збору даних.

Центральний елемент модуля навчання – скрипт train.py, що реалізує логіку навчання, порівняння різних видів машинного навчання та абляційний аналіз. Вхідні дані зберігаються у файлі FULL_DATA.csv, сформованому консольною утилітою збору даних на основі консолідації записів FAOSTAT та агрегованих

кліматичних індикаторів Open-Meteo. Структуру скрипту побудовано за модульним принципом і містить реалізації трьох незалежних алгоритмів: лінійної регресії (`linearregressionlearn`), випадкового лісу (`rflearn`) та градієнтного бустингу (`xboostlearn`), що дозволяє виконувати порівняльний аналіз їхньої ефективності в межах єдиного експериментального середовища.

Першим етапом конвеєра обробки даних є попереднє оброблення для усунення шуму та формування вхідного вектора ознак. Щоб мінімізувати вплив аномальних викидів, характерних для агростатистичних даних (наприклад, помилок обліку або катастрофічних неврожаїв), застосовано метод квантильної фільтрації. Алгоритм відсіює спостереження, що виходять за межі діапазону $[0,01; 0,99]$ розподілу цільової змінної `YieldGot`, забезпечуючи навчання моделі на репрезентативній вибірці.

Для програмної реалізації розрахунку показника технологічного рівня країни (CTL), що визначено як середнє математичне сподівання врожайності, використано методи бібліотеки `Pandas` [75]. Замість ітеративних обчислень застосовано векторизовану операцію групування та трансформації, що забезпечує високу швидкість обробки великих масивів даних.

Нормалізацію розподілу цільової змінної (урожайності) реалізовано через логарифмічну трансформацію $\ln(1+Y)$, що дає нівелювати вплив гетероскедастичності даних. У бібліотеці `NumPy` для цього використано функцію `log1p`, котра забезпечує вищу точність для малих значень аргументу порівняно зі звичайним логарифмом [76].

Щоб побудувати кінцеву прогностну модель, використано алгоритм `XGBRegressor` з бібліотеки `XGBoost`. Конфігурацію гіперпараметрів (кількість естиматорів $n=500$, швидкість навчання $\eta=0.05$, максимальна глибина дерева $d=6$) обрано для досягнення балансу між здатністю моделі відтворювати складні нелінійні залежності та запобіганням перенавчанню. Валідація результатів здійснюється через розбиття вибірки на тренувальну та тестову в пропорції 80/20,

а також із застосуванням 5-блокової перехресної перевірки, що забезпечує статистичну достовірність отриманих метрик якості [77].

Важливим етапом роботи пайплайну є не лише збереження ваг навченої моделі, а й обчислення коригувального коефіцієнта для переходу від логарифмічного прогнозу до реальних фізичних величин. Програмний модуль обчислює дисперсію залишків на тестовій вибірці (`sigma_squared`), що використовується для корекції зміщення оцінки математичного сподівання.

Результатом роботи конвеєра навчання є серіалізований файл моделі *model.bin*, що містить оптимізовані ваги ансамблю дерев. Цей артефакт є відчуженим від Python-середовища і надалі інтегрується в клієнтський додаток як статичний ресурс, що уможливило прогнозування без розгортання інтерпретатора Python на машині кінцевого користувача.

3.1.3 Структура runtime додатку та модулів бізнес-логіки. Програмне рішення реалізовано у вигляді єдиного комплексу YieldPredictor, архітектурна декомпозиція якого виконана відповідно до принципів багатоварової архітектури [78]. Такий підхід забезпечив логічне розмежування компонентів, що відповідають за бізнес-правила, інтерфейс користувача та взаємодію із зовнішніми джерелами даних, гарантуючи слабку зв'язність модулів та високу придатність до масштабування.

Функціональним ядром системи є бібліотека класів YieldPredictor.Core. Цей модуль концентрує в собі всю предметну логіку та не залежить від деталей реалізації інтерфейсу користувача. У його структурі виділено шар доменних моделей, що описують сутності агрономічної предметної області (кліматичні дані, характеристики добрив, результати прогнозування), та шар сервісів, що реалізують алгоритми обробки інформації. Саме тут інкапсульовано механізми розрахунку агрокліматичних індексів, логіку перетворення фізичної ваги добрив у діючу речовину та алгоритми оптимізації за прибутковістю. Така ізоляція ядра дозволяє використовувати розроблені алгоритми в різних типах застосунків (веб, десктоп, мобільні) без потреби модифікувати програмний код.

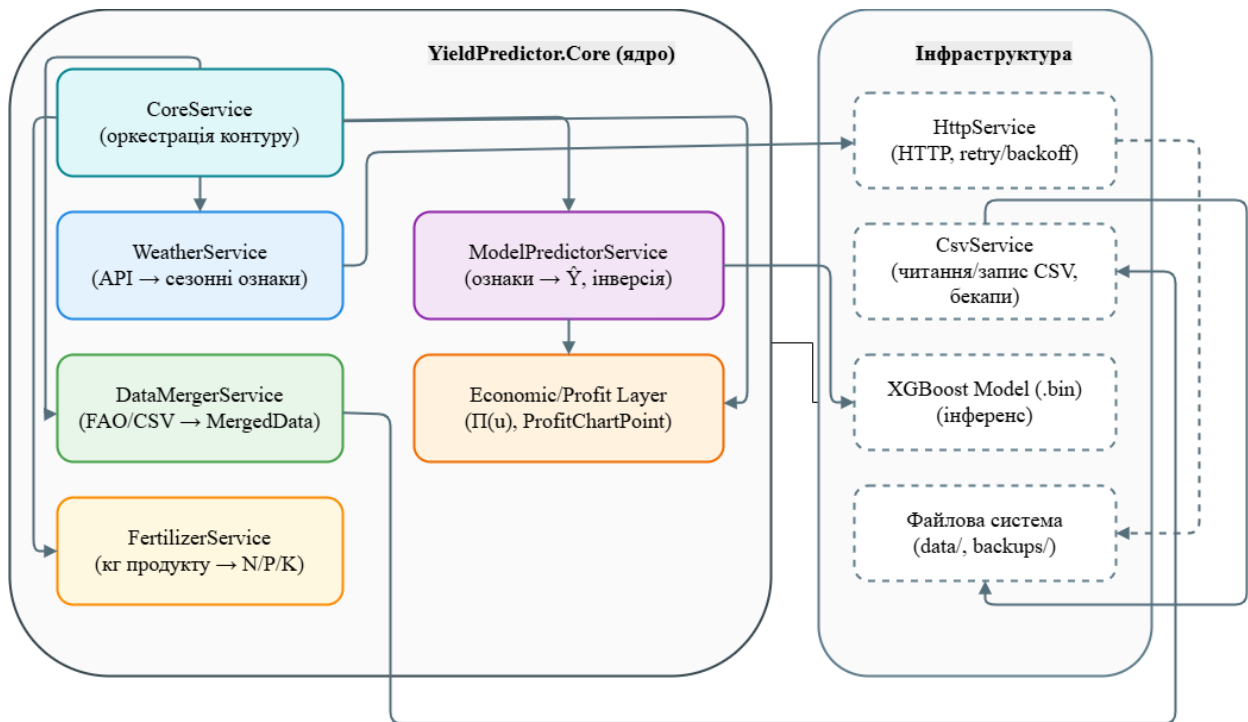


Рис. 3.2. Діаграма компонентів програмного комплексу YieldPredictor

Підсистему взаємодії з користувачем реалізовано в модулі YieldPredictor.Client. Цей компонент виконує роль презентаційного шару та відповідає за ініціалізацію системи, валідацію вхідних даних та візуалізацію результатів роботи аналітичного ядра. Внутрішню структуру модуля побудовано на базі XAML-розмітки, що визначає візуальне компонування вікон, та пов'язаного коду, який обробляє події інтерфейсу. Модуль забезпечує асинхронний виклик методів ядра, що гарантує відгук інтерфейсу навіть під час виконання тривалих обчислювальних операцій.

Для забезпечення комунікації із зовнішнім середовищем у структурі комплексу виділено інфраструктурні сервіси. Вони відповідають за низькорівневу реалізацію протоколів обміну даними, зокрема виконання HTTP-запитів до API кліматичних сервісів та геокодування. Ці компоненти реалізують механізми оброблення мережових помилок, повторних запитів та кешування даних, абстрагуючи бізнес-логіку від технічних деталей мережової взаємодії. Інтеграцію з файловою системою забезпечують спеціалізовані сервіси

серіалізації, що дозволяють зберігати та відновлювати стан системи, а також експортувати результати розрахунків у форматі CSV.

Окремим структурним елементом є модуль інтеграції машинного навчання, реалізований через адаптер `ModelPredictorService`. Цей компонент забезпечує завантаження попередньо навченої моделі `XGBoost` із бінарного файлу та виконання інференсу в середовищі `.NET`. Використання бібліотеки-обгортки `XGBoostSharp` дало змогу інтегрувати високопродуктивний `C++` код бібліотеки бустингу безпосередньо в керований код `C#`, забезпечивши передачу даних між об'єктною моделлю застосунку та масивами ознак, потрібними для роботи ансамблевої моделі `XGBoost`.

Візуалізацію аналітичних даних здійснює інтегрований компонент побудови графіків на основі бібліотеки `ScottPlot`. Цей модуль відповідає за рендеринг інтерактивних діаграм прибутковості, автоматичне масштабування осей та відображення точок оптимуму. Архітектурно його відокремлено від розрахункового ядра: ядро лише формує масиви числових даних, тоді як модуль візуалізації відповідає тільки за їх графічне представлення, що відповідає принципу єдиної відповідальності.

3.1.4 Взаємодія модулів. Забезпечення надійності, розширюваності та придатності до тестування програмного комплексу досягнуто через імплементацію базових принципів об'єктно-орієнтованого проєктування, зокрема принципів `SOLID` [79]. Архітектура системи ґрунтується на слабкій зв'язності компонентів, що реалізовано через низку породжувальних та структурних патернів проєктування.

Для керування життєвим циклом об'єктів та приховування складності ініціалізації базових компонентів застосовано фабричний метод [80]. Логіку створення головного сервісу ядра інкапсульовано у спеціалізованому класі `CoreServiceFactory`. Таке архітектурне рішення звільняє клієнтський рівень застосунку від потреби в ручному конструюванні залежностей, таких як клієнти мережевої взаємодії чи парсери файлів. Натомість презентаційний шар

звертається до фабрики й отримує повністю налаштований і готовий до роботи екземпляр ядра, що уможлиблює централізоване керування конфігураціями шляхів до моделей машинного навчання та локальних баз даних.

Організацію зв'язків між внутрішніми підсистемами побудовано на основі патерну впровадження залежностей. Головний клас-оркестратор `CoreService` отримує всі потрібні для функціонування компоненти через конструктор у вигляді абстракцій. Такий підхід забезпечує інверсію керування [81; 82]: ядро системи не залежить від конкретних реалізацій мережесих клієнтів чи предикторів, а покладається тільки на їхні інтерфейси. Це важлива умова для відтворюваності обчислювальної апробації та верифікації моделі, адже це дає змогу під час тестування ізолювати бізнес-логіку та підміняти реальні зовнішні сервіси імітаційними об'єктами.

Комунікацію між модулями регламентує принцип розділення інтерфейсів. Взаємодія здійснюється через чітко визначені контракти, такі як `IWeatherService`, `IFertilizerService` та `IModelPredictorService`. Наприклад, сервіс оркестрації оперує винятково сигнатурою методу прогнозування та не має доступу до внутрішньої логіки оброблення тензорів або специфіки роботи C++ бібліотеки `XGBoost`, яка прихована на рівні реалізації. Таке абстрагування мінімізує ризики каскадних помилок у разі оновлення або заміни алгоритмічного апарату.

Основний функціональний процес системи – побудова простору економічних сценаріїв для визначення оптимальної дози добрив реалізується через чітку послідовність оркестрації підпорядкованих сервісів. Алгоритм функціонує як цілісний конвеєр обробки даних. На початковому етапі виконується ініціалізація обчислювального контексту: система отримує географічні координати цільової ділянки та ідентифікатор обраного типу добрива. Далі ініціюється збір природно-кліматичних та технологічних даних: виконується асинхронний запит для агрегації погодних умов за вегетаційний період та паралельно визначається макроекономічний технологічний рівень регіону.

Після формування контексту алгоритм ініціює циклічну симуляцію простору рішень, застосовуючи метод перебору по сітці для набору допустимих норм внесення добрив. Для кожної ітерації виконується конвертація фізичної маси препарату у стандартизований вектор діючих речовин (азот, фосфор, калій). Сформований інтегрований вектор ознак, що об'єднує кліматичні, агрохімічні та технологічні параметри, передається до модуля машинного навчання, що генерує прогноз очікуваної врожайності. На завершальному кроці ітерації система здійснює економічну оцінку: зіставляє прогнозовану виручку від реалізації врожаю з витратами на внесену дозу добрив, обчислюючи чистий прибуток. Згенерований масив сценарних точок формує криву прибутковості, що повертається на презентаційний шар для візуалізації та прийняття кінцевого управлінського рішення.

З огляду на високу залежність обчислень від операцій введення-виведення та мережевої взаємодії, системну архітектуру побудовано на засадах асинхронної моделі опрацювання даних. Усі процеси, пов'язані із зовнішніми запитами або зверненнями до файлової системи, реалізовано з використанням механізмів неблокуючого виконання (патерн `async/await`). Таке архітектурне рішення унеможливорює зависання головного потоку застосунку [83] під час очікування відповідей від віддалених серверів, що забезпечує високу чутливість інтерфейсу та стабільну пропускну здатність системи навіть за умов опрацювання великих масивів історичних даних.

3.1.5 Конфігурація системи. Забезпечення гнучкості й адаптивності розробленого програмного комплексу до змінних умов експлуатації вимагає впровадження ефективних механізмів керування параметрами середовища. Для цього в системі реалізовано принцип екстерналізації конфігурацій. Замість жорсткого кодування системних змінних безпосередньо в програмному коді, управління параметрами винесено у спеціалізовані конфігураційні файли формату JSON [84]. Такий підхід дозволяє динамічно змінювати ключові налаштування системи – зокрема, шляхи до серіалізованих моделей машинного

навчання, адреси кінцевих точок зовнішніх API та параметри тайм-аутів мережеских з'єднань – без потреби перекомпільовувати весь проєкт. Це суттєво спрощує розгортання застосунку на кінцевих робочих станціях та полегшує його супровід.

3.1.6 Інтерфейс користувача. Підсистему людино-машинної взаємодії (графічний інтерфейс користувача) спроектовано на базі платформи Windows Presentation Foundation (WPF). Використання декларативної мови розмітки XAML забезпечило чітке архітектурне розмежування між візуальним поданням інформації та керівною програмною логікою. Такий розподіл відповідальності гарантує, що модифікація візуальних елементів (наприклад, зміна структури форм вводу чи кольорових схем) жодним чином не впливає на цілісність аналітичного ядра системи. Крім того, застосування вбудованих механізмів прив'язки даних дало змогу автоматизувати синхронізацію стану внутрішніх моделей об'єктів із відповідними елементами інтерфейсу.

Для забезпечення роботи з просторовими даними в середовищі WPF інтегровано компонент WebView2, що функціонує на основі рушія Chromium [85]. Це рішення дало змогу ізолювати геоінформаційну логіку (роботу з картографічними сервісами) у вебконтейнері, забезпечивши при цьому захищений і високопродуктивний асинхронний обмін повідомленнями між інтерфейсом користувача та аналітичним ядром системи.

Процес взаємодії користувача з аналітичним ядром реалізовано з урахуванням принципів стійкості до помилок введення. На етапі роботи клієнтського застосунку виконується попередня валідація даних: перед запуском конвеєра обробки система автоматично перевіряє введені користувачем географічні координати, цінові параметри та обрані типи добрив на відповідність допустимим фізичним і економічним діапазнам. У разі виявлення аномалій (наприклад, введення від'ємної вартості ресурсу або неіснуючих координат) алгоритм блокує виконання обчислень і генерує відповідне інформаційне

сповіщення, запобігаючи нераціональному використанню обчислювальних ресурсів та формуванню хибних прогнозів.

Щоб забезпечити зручність застосунку та безперервний зворотний зв'язок, усі звернення від інтерфейсу до обчислювального ядра виконуються виключно в асинхронному режимі. Це означає, що під час виконання тривалих мережових запитів або складних матричних обчислень у модулі машинного навчання графічна оболонка залишається чутливою до дій користувача, супроводжуючи процес індикаторами прогресу.

Для інтерпретації результатів моделювання в інтерфейс інтегровано модуль візуалізації на базі високоефективної графічної бібліотеки ScottPlot. Цей компонент відповідає за динамічний рендеринг кривих прибутковості та функцій відгуку врожайності в режимі реального часу. Перетворення масивів числових даних на наочне графічне подання дозволяє особі, що приймає рішення, візуально ідентифікувати точку глобального економічного оптимуму й оцінити межі доцільності збільшення норм внесення мінеральних добрив, що є кінцевою метою функціонування всієї інформаційної системи.

3.2. Алгоритмічне забезпечення агрономічного, прогнозного та економічного компонентів

3.2.1 Інформаційне забезпечення та інтеграція зовнішніх джерел даних.

Функціонування інтелектуальної системи підтримки аграрних рішень прямо залежить від якості, повноти та репрезентативності вхідних інформаційних масивів. З огляду на глобальний масштаб поставленої задачі та потребу врахування макроекономічних, технологічних і природно-кліматичних чинників, інформаційне забезпечення розробленого програмного комплексу побудовано на принципах інтеграції відкритих міжнародних баз даних та спеціалізованих геоінформаційних вебсервісів. Такий підхід забезпечує незалежність системи від локальних платних джерел та гарантує відтворюваність результатів моделювання в умовах географічної мінливості.

Основним джерелом агростатистичної інформації для формування навчальної вибірки та визначення базових технологічних показників слугує база даних Продовольчої та сільськогосподарської організації ООН (FAOSTAT). Цей ресурс надає доступ до багаторічних історичних рядів щодо середньої врожайності сільськогосподарських культур та обсягів використання мінеральних добрив у розрізі країн світу. Використання масивів FAOSTAT дозволяє сформувати глобальну репрезентативну вибірку, що охоплює різноманітні технологічні уклади та ґрунтово-кліматичні зони. З алгоритмічної погляду, структура цих даних потребує попередньої трансформації: обсяги споживання добрив перераховуються у базові елементи живлення – азот, фосфор та калій (NPK) – за часткою діючої речовини. Цей етап є важливою умовою забезпечення адекватності алгоритмів оптимізації, адже механічне агрегування різних типів добрив в єдину фізичну масу призвело б до втрати інформації про обмежувальні фактори росту та порушення базових законів агрохімії.

Щоб забезпечити прогнозну модель ретроспективними й оперативними кліматичними даними в архітектуру системи, інтегровано програмний інтерфейс Open-Meteo Archive API. Вибір цього сервісу зумовлено доступом до глобальних масивів метеорологічного реаналізу, що усуває жорстку прив'язаність програмного комплексу до конкретних наземних метеостанцій, які нерідко допускають пропуски у спостереженнях або мають недостатє просторове покриття. Через механізми API система отримує добові значення максимальної та середньої температур, а також суми опадів для будь-якої заданої географічної координати. На наступному етапі алгоритми обробки здійснюють фільтрацію неактивних періодів року й агрегують отримані добові показники у спеціалізовані сезонні індикатори. До таких індикаторів належать суми ефективних температур та індекси накопиченого теплового стресу, логіка розрахунку яких базується на теоретичних біологічних моделях температурних мінімумів та максимумів, описаних у попередньому розділі.

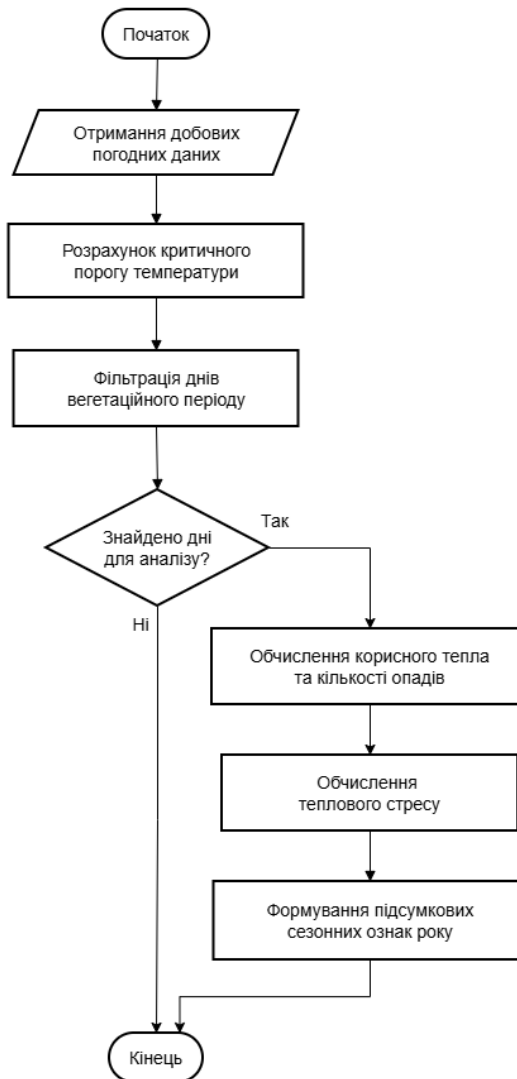


Рис. 3.3. Блок-схема агрегації погодних рядів для країни

Для просторової синхронізації локальних запитів користувача із макрорегіональними статистичними показниками застосовано технологію зворотного геокодування. Інформаційним донором для цього виступає відкритий сервіс Nominatim, що функціонує на базі картографічних даних OpenStreetMap. Алгоритм взаємодії передбачає відправлення мережевого запиту із заданими користувачем широтою та довготою ділянки, після чого сервіс повертає структуровану інформацію про адміністративну приналежність об'єкта, зокрема назву держави. Система використовує отриману геолокаційну інформацію для автоматичного зіставлення із внутрішньою базою знань, щоб вилучити відповідний технологічний індекс країни. Цей показник обчислено як середня

врожайність по країні на тренувальній вибірці, що слугує статистичним проксі загального технологічного рівня агровиробництва в обраному макрореґіоні.

3.2.2 Алгоритми машинного навчання для прогнозування врожайності.

Центральне завдання аналітичного ядра розробленої системи – апроксимація складної нелінійної функції відгуку врожайності на внесення мінеральних добрив за умов специфічного кліматичного фону. Щоб розв'язати цю задачу та вибрати оптимальний прогностичний апарат на етапі дослідження, імплементовано та досліджено декілька алгоритмів машинного навчання, що суттєво різняться за своєю архітектурою, обчислювальною складністю та здатністю узагальнювати просторові закономірності.

Як базовий еталонний метод для оцінки ефективності складніших моделей використано алгоритм множинної лінійної регресії. Цей підхід ґрунтується на мінімізації сум квадратів відхилень між прогнозованими та фактичними значеннями, формуючи оптимальну гіперплощину в багатовимірному просторі ознак. Безперечна перевага лінійної моделі – швидкодія та прозорість впливу кожного чинника, що виражається через вагові коефіцієнти при незалежних змінних. Проте, з огляду на специфіку предметної області, алгоритм спирається на припущення про лінійність і адитивність зв'язків. Відповідно, він не здатен адекватно відтворити синергетичні ефекти (наприклад, взаємодію між дозою внесеного азоту та кількістю опадів) або ефекти фізіологічного насичення, коли подальше збільшення норми добрив припиняє давати приріст врожаю і навіть призводить до його зниження (токсичний ефект). Тому лінійну регресію визнано недостатньо гнучкою для кінцевого використання в системі підтримки рішень.

Щоб подолати обмеження лінійних підходів та врахувати нелінійну взаємодію кліматичних і технологічних чинників, застосовано алгоритм випадкового лісу [86; 87]. Цей алгоритм належить до класу ансамблевих методів і базується на концепції беггінгу (bagging – bootstrap aggregating). Процес навчання полягає в побудові великої кількості незалежних дерев рішень, кожне з яких навчається на випадковій підмножині даних із поверненням, а для

розгалуження у вузлах використовується лише випадкова підмножина ознак. Це забезпечує декореляцію базових моделей та суттєво знижує дисперсію фінального прогнозу, що формується через усереднення результатів усіх дерев. Випадковий ліс продемонстрував високу стійкість до аномальних викидів у сільськогосподарській статистиці. Попри це, виявлено суттєві недоліки методу в контексті програмної реалізації: збереження ансамблю із сотень глибоких дерев потребує значних обсягів оперативної пам'яті, а швидкість виконання інференсу (отримання прогнозу) є порівняно низькою, що може перешкодити швидкому перебору сценаріїв в інтерактивному десктопному застосунку.

Оптимальним алгоритмічним рішенням, яку обрано для фінальної інтеграції в клієнтську частину програмного комплексу, став алгоритм екстремального градієнтного бустингу. На відміну від випадкового лісу, де дерева будуються паралельно та незалежно, градієнтний бустинг використовує секвенційну архітектуру: кожне наступне дерево цілеспрямовано фокусується на виправленні помилок, допущених сумою всіх попередніх дерев в ансамблі.

Ключова математична особливість алгоритму XGBoost [66] полягає в тому, що він виконує оптимізацію цільової функції втрат у функціональному просторі, використовуючи наближення функції Тейлора до другого порядку. На відміну від класичного градієнтного бустингу, який спирається лише на першу похідну (градієнт), XGBoost додатково обчислює другу похідну (гессіан). Це забезпечує значно швидшу й точнішу збіжність алгоритму до глобального мінімуму.

Крім того, в архітектуру XGBoost на рівні цільової функції вбудовано механізми L1 (Lasso) та L2 (Ridge) регуляризації ваг листків дерев [88; 89]. Цей інструментарій жорстко штрафує модель за надмірну структурну складність, тим самим ефективно запобігаючи перенавчанню на інформаційних шумах, які є іманентною властивістю будь-яких агростатистичних даних. Додатковою перевагою алгоритму є його здатність автоматично обробляти розріджені матриці та дані з пропусками, скеровуючи такі спостереження в гілки дерев, що мінімізують загальну похибку.

Висока точність апроксимації нелінійних функцій відгуку разом з апаратною оптимізацією структур даних у пам'яті забезпечила алгоритму XGBoost відповідно високі показники швидкості інференсу. Це дало змогу інтегрувати навчену модель безпосередньо в середовище .NET-застосунку, забезпечивши миттєвий розрахунок тисяч можливих агротехнологічних сценаріїв у режимі реального часу.

3.2.3 Алгоритми попередньої обробки даних та оптимізації за прибутковістю з урахуванням ризику. Ефективність функціонування предиктивних моделей машинного навчання залежить від якості вхідного простору ознак. Відповідно, процес підготовки даних до подачі в алгоритм XGBoost та подальша інтерпретація отриманих результатів вимагають застосування спеціалізованого комплексу алгоритмів попередньої та постобробки. З огляду на специфіку розподілу сільськогосподарських показників, алгоритмічне забезпечення системи передбачає обов'язкову нормалізацію цільової змінної.

Статистичний аналіз історичних даних засвідчив, що показники врожайності характеризуються вираженим асиметричним розподілом та мультиплікативною природою похибок. Щоб стабілізувати дисперсію та привести розподіл до нормального вигляду на етапі навчання, застосовано алгоритм логарифмічного перетворення цільової змінної. З математичного погляду це означає, що градієнтний бустинг оптимізується не на абсолютних, а на відносних похибках прогнозу.



Рис. 3.4. Блок-схема точкового прогнозу врожайності

Під час виконання інференсу в середовищі клієнтського застосунку виникає зворотна задача – трансформація логарифмічного прогнозу в реальні фізичні величини (кілограми на гектар). Для цього застосовується алгоритм експоненціювання. Проте математичні властивості логнормального розподілу зумовлюють виникнення систематичного зміщення. Щоб нівелювати цей ефект, в алгоритм постобробки імплементовано механізм аналітичної корекції, який використовує заздалегідь обчислену на тестовій вибірці дисперсію залишків як коригувальний коефіцієнт. Такий підхід гарантує отримання математично незміщеної й економічно адекватної оцінки планової врожайності.

Після отримання валідованого агрономічного прогнозу система переходить до розв'язання задачі пошуку економічно доцільної норми внесення добрив. Для цього розроблено алгоритм дискретної оптимізації методом просторового сценарного аналізу. Сформована ансамблем дерев функція відгуку є суворо нелінійною, багатовимірною та не має явного аналітичного виразу для обчислення градієнта (похідної), тому неможливо застосувати класичні методи математичного аналізу для знаходження екстремуму. Натомість реалізовано детермінований ітеративний алгоритм прямого пошуку.

Процедура оптимізації ініціюється шляхом генерації одновимірної дискретної сітки допустимих значень агрохімічних норм. Діапазон сітки та крок ітерації формуються з урахуванням фізіологічно виправданих меж та технологічних обмежень сільськогосподарської техніки (дозаторів розкидачів добрив). Для кожної згенерованої точки (сценарію) алгоритм ініціює виклик прогнозової моделі, обчислюючи очікувану врожайність за фіксованих поточних кліматичних умов.

На наступному етапі алгоритм накладає на отриманий масив агрономічних прогнозів актуальні макроекономічні параметри: прогнозовану ринкову ціну реалізації сільськогосподарської культури та вартість конкретного типу мінерального добрива. Для кожного сценарію обчислюється функція маржинального прибутку, що визначається як різниця між очікуваною виручкою та прямими витратами на агрохімікати. З отриманого масиву пар «Доза – Прибуток» алгоритм ідентифікує той вектор, що забезпечує глобальну максимізацію чистого фінансового результату.

Такий підхід гарантує знаходження абсолютного оптимуму в заданому просторі технологічних рішень. Крім того, збереження результатів усіх проміжних ітерацій дає змогу побудувати неперервну криву економічної ефективності, що передається на клієнтський рівень системи для візуального аналізу та прийняття остаточного управлінського рішення користувачем.

Окрім детермінованого пошуку оптимального рішення за фіксованих умов, базовий алгоритм сценарного аналізу розширено модулем оцінки кліматичних ризиків. З огляду на стохастичну природу погодних умов, прийняття рішень суто на основі усереднених багаторічних показників часто призводить до формування надто оптимістичних та фінансово вразливих стратегій. Щоб розв'язати цю проблему, розроблено алгоритм оптимізації внесення добрив за прибутковістю з урахуванням ризику, який базується на принципах сценарної генерації (метод історичного моделювання Монте-Карло).

Логіка ризик-орієнтованого алгоритму полягає у формуванні багатовимірному простору результатів. Замість використання єдиного вектора середніх температур та опадів алгоритм завантажує колекцію ретроспективних погодних сценаріїв за певний проміжок часу (наприклад, останні 10–20 років). Для кожної фіксованої дози добрив із дискретної сітки система виконує множинні виклики прогнозної моделі та послідовно підставляє параметри кожного історичного року.

У результаті для кожної норми внесення агрохімікатів алгоритм генерує не єдине точкове значення, а повноцінний масив очікуваних врожайностей, що згодом трансформується у "віяло" ймовірних фінансових результатів (розподіл прибутків). Щоб оцінити міру ризику, розроблений алгоритм сортує отриманий масив прибутків за зростанням. Далі система автоматично виокремлює нижній хвіст розподілу – заданий квантиль найгірших фінансових сценаріїв (наприклад, 10% найзбитковіших років, які відповідають умовам жорсткої посухи).

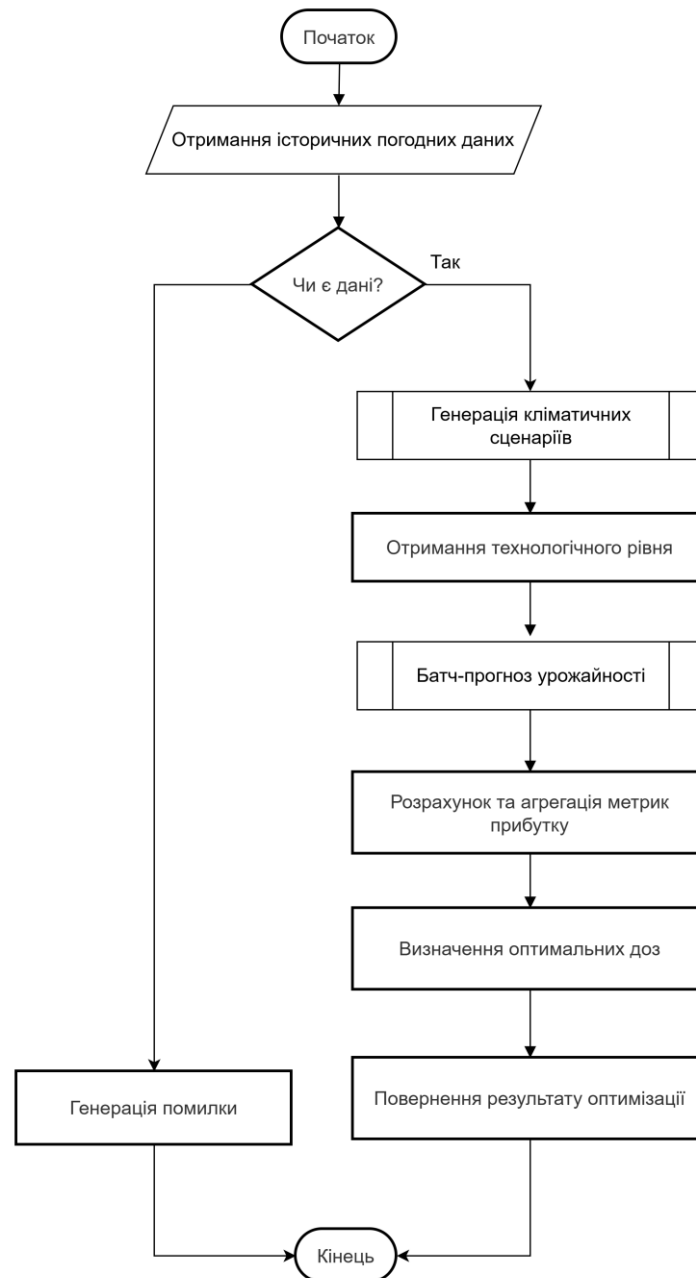


Рис. 3.5. Блок-схема оптимізації за прибутковістю з урахуванням ризику

Показник умовної вартості під ризиком (CVaR), що кількісно відображає середній очікуваний прибуток у найгірших кліматичних ситуаціях; від'ємне значення інтерпретується як очікуваний збиток у цих сценаріях. Кінцеве управлінське рішення ідентифікується алгоритмом не через пошук абсолютного максимуму прибутку, а через максимізацію збалансованої цільової функції, що враховує штраф за ризик. Такий підхід дає системі змогу визначати

"найбезпечнішу" дозу добрив, яка гарантує збереження фінансової стійкості агропідприємства навіть за настання несприятливих погодних умов.

3.3. Реалізація прототипу інтелектуальної системи на базі інтегрованих даних про погоду

3.3.1 Реалізація інтеграції з зовнішніми API та джерелами даних. Процес програмної імплементації прототипу інтелектуальної системи розпочато зі створення інфраструктурного шару, головне завдання якого – забезпечити надійний збір, парсинг і попередню агрегацію інформації з розрізнених зовнішніх джерел. Архітектурно цей шар ізолює ядро бізнес-логіки від низькорівневих операцій вводу-виводу та деталей мережевої взаємодії, що відповідає принципам чистої архітектури.

Базовим компонентом мережевої взаємодії виступає спеціалізований клас `HttpService`, що функціонує як обгортка над стандартним класом `HttpClient` платформи .NET. Використання єдиного статичного або інжектованого екземпляра клієнта в межах усього життєвого циклу застосунку дозволило уникнути проблеми вичерпання мережевих сокетів, типової при інтенсивному потоковому завантаженні великих масивів метеоданих. Програмна реалізація `HttpService` використовує вбудовані механізми `System.Text.Json` для ефективної десеріалізації відповідей у строго типізовані об'єкти (зокрема через дженерик-метод `GetAsync<T>`), що мінімізує витрати оперативної пам'яті на парсинг великих JSON-документів.

Отримання ретроспективних кліматичних показників делеговано класу `WeatherService`, який імплементує логіку взаємодії з відкритим програмним інтерфейсом `Open-Meteo Archive API`. Алгоритм формує параметризовані HTTP-запити, динамічно підставляючи координати (широту та довготу), заданий часовий діапазон та перелік необхідних метеорологічних змінних (середня і максимальна температури, сума опадів). З огляду на жорсткі обмеження відкритих API на кількість запитів за хвилину, у програмному коді цього сервісу

реалізовано важливий механізм відмовостійкості на базі патерну повторних спроб із затримкою.

Програмно цей механізм вирішено через імплементацію циклу `while`, усередині якого виконується мережевий виклик із перехопленням винятків типу `HttpRequestException`. У разі отримання від сервера статус-коду 429 (`Too Many Requests`) потік виконання не переривається аварійно, а призупиняється: реалізовано фіксовану паузу 60 с при HTTP 429; при інших мережевих помилках застосовується лінійне збільшення затримки.

Щоб забезпечити збереження прогресу при пакетній обробці сотень локацій, `WeatherService` додатково імплементує механізм проміжного резервного копіювання. Після успішного завантаження та агрегації порції даних система серіалізує їх у локальний файл, що дозволяє відновити процес збирання інформації з точки збою без необхідності повторного звернення до API.

Операції з локальною файловою системою, зокрема читання завантажених історичних баз даних агростатистики (`FAOSTAT`), реалізовано в класі `CsvService` за допомогою оптимізованої бібліотеки `CsvHelper`. Цей компонент ініціалізується з налаштуваннями `CultureInfo.InvariantCulture` для коректного парсингу числових значень із глобальних таблиць.

Завершальним етапом інфраструктурної підготовки є консолідація розрізнених потоків інформації у єдиний простір ознак, що реалізовано в класі `DataMergerService`. Алгоритм злиття розроблено з інтенсивним використанням технології інтегрованих запитів (`LINQ`). На першому етапі масиви даних про врожайність та внесення добрив групуються за складеним ключем, що включає ідентифікатор країни та рік спостереження (`GroupBy(x => new { x.Country, x.Year })`). Під час обробки інформації стосовно добрив система виконує семантичний аналіз назв показників (через метод `IndexOf(..., StringComparison.OrdinalIgnoreCase)`), автоматично виокремлюючи та підсумовуючи обсяги споживання азоту, фосфору та калію. Після агрегації виконується операція внутрішнього з'єднання (`Inner Join`) колекцій, у результаті

якої формується фінальний строго типізований масив об'єктів `MergedData`, готовий до збагачення погодними показниками та подачі на вхід моделі машинного навчання.

Технічна реалізація контекстної ознаки `CountryTechLevel` винесена в окремий компонент `CtlCalculator`, що реалізує інтерфейс `ICtlCalculator`. Розрахунок виконується одноразово на основі повного історичного масиву даних FAOSTAT за обраним переліком продовольчих культур, незалежно від подальшого розбиття вибірки на навчальну й тестову.

Алгоритм обчислення складається з трьох етапів. На першому етапі для кожної культури з переліку обчислюється середньосвітове значення врожайності шляхом групування записів за назвою культури (`Item`) і усереднення показника `Value` по всіх країнах. На другому етапі дані групуються за країною (`Area`), і для кожної культури, представленої в записах конкретної країни, обчислюється відношення середньої врожайності цієї культури в країні до її середньосвітового значення. На третьому етапі отримані по кожній культурі коефіцієнти усереднюються в межах країни, формуючи підсумкове значення `CountryTechLevel`.

Реалізація передбачає обробку записів із пропущеними значеннями (`Value.HasValue`) шляхом їх попереднього вилучення з вибірки до початку розрахунків, що виключає викривлення середніх значень через відсутні дані. Якщо для певної культури середньосвітове значення дорівнює нулю (теоретично можливий, хоча практично малоймовірний випадок), коефіцієнт для цієї культури в межах країни приймається рівним нулю, що запобігає діленню на нуль.

Розраховані коефіцієнти серіалізуються у формат JSON у вигляді відображення «країна → коефіцієнт» і завантажуються на етапі формування вектора ознак для кожного спостереження через зіставлення за назвою країни. Для країн, відсутніх у словнику коефіцієнтів (через брак даних FAOSTAT за обраним переліком культур), застосовується резервне значення, що дорівнює середньому коефіцієнту по всіх країнах словника

3.3.2. Реалізація очищення даних. Використання глобальних масивів відкритої статистики (FAOSTAT) та метеорологічного реаналізу (Open-Meteo) об'єктивно супроводжується проблемою якості первинних даних. Пропущені значення, помилки введення та методологічні невідповідності у звітності різних країн потребують застосування ефективної процедури попереднього опрацювання, в іншому разі інтеграція таких даних в аналітичне ядро призводить до зниження точності прогнозування моделі машинного навчання. Для формування однорідного простору ознак у системі реалізовано програмний модуль ідентифікації й усунення інформаційних аномалій.

Перший етап очищення спрямовано на опрацюванні масивів агрохімічної та статистичної інформації. Специфіка зведених глобальних баз даних полягає в тому, що окремі держави можуть періодично змінювати стандарти звітності, наприклад, вказуючи обсяги внесених добрив у тоннах замість кілограмів, або фіксувати врожайність у центнерах замість грамів на квадратний метр. Такі розбіжності формують у вибірці аномальні викиди, здатні суттєво спотворити результати навчання моделі під час навчання моделі.

Щоб розв'язати цю проблему, імплементовано алгоритм квантильної фільтрації простору ознак. Замість ручного видалення підозрілих записів система автоматично обчислює статистичний розподіл цільової змінної (врожайності) та предикторів (доз добрив) для кожного макрорегіону. Алгоритм відсікає спостереження, що виходять за межі діапазону від 1-го до 99-го перцентиля. Цей підхід ефективно ізолює очевидні технічні помилки звітності та вплив нетипових екстремальних явищ, залишаючи для навчання моделі тільки репрезентативні дані, що відображають нормальні агрономічні процеси.

Другий етап очищення стосується оброблення пропущених значень (NaN) у часових рядах метеорологічних показників. Попри високу надійність API-сервісів реаналізу, у завантажених масивах періодично трапляються пропуски. Оскільки алгоритм розрахунку сезонних індикаторів потребує достовірних даних, рядки з відсутніми значеннями кліматичних показників виключаються з вибірки

на етапі підготовки даних. Такий підхід гарантує безперебійну роботу математичного апарату системи, формуючи чистий та достовірний масив предикторів для алгоритму XGBoost.

3.3.3 Реалізація обробки та інтеграції даних. Після успішного отримання та первинної консолідації необроблених даних від зовнішніх сервісів ключовим завданням інформаційної системи є семантичне перетворення цих масивів у спеціалізовані предметні показники (видобування ознак). Програмну реалізацію цього етапу розділено на два незалежні функціональні блоки: підсистему агрокліматичної агрегації та підсистему агрохімічних розрахунків, що забезпечує дотримання принципу єдиної відповідальності при проєктуванні класів.

Первинне опрацювання метеорологічних часових рядів реалізовано в межах класу WeatherService. Зовнішній API повертає дані у вигляді масивів добових значень за весь рік, тож система повинна виокремити виключно період активної вегетації. Щоб оптимізувати використання оперативної пам'яті під час парсингу відповідей сервера, замість повної об'єктної десеріалізації застосовано низькорівневий підхід на базі структури JsonDocument із простору імен System.Text.Json [92]. Це дає змогу виконувати пряму навігацію по DOM-дереву JSON-відповіді через методи GetProperty та EnumerateArray та вилучати лише необхідні масиви температур і опадів без створення надлишкових проміжних об'єктів у пам'яті.

Програмна реалізація алгоритму розрахунку суми ефективних температур (GDD) та індексу теплового стресу (SDD) передбачає однопрохідне ітеративне обходження синхронізованих добових масивів. Усередині циклу реалізовано логіку фільтрації: якщо середня добова температура перевищує встановлений біологічний мінімум (константа базової температури 5°C), ця різниця акумулюється в загальну суму GDD. Дні з температурою, нижчою за цей поріг, ігноруються алгоритмом за допомогою умовного оператора continue, що дозволяє оптимізувати проходження циклу. Одночасно в цьому самому циклі

підраховується сумарна кількість опадів за вегетаційний період, що мінімізує обчислювальну складність алгоритму до $O(n)$, де n – кількість днів у році.

Особливістю програмної реалізації розрахунку теплового стресу (SDD) є використання динамічного порогу, що адаптується до географічної широти місцевості. Граничні значення спираються на агрономічні дослідження теплового стресу: поріг $\sim 30\text{--}31^\circ\text{C}$ для культур помірних широт підтверджений як класичними оглядами (Porter & Gawith, 1999) [103], так і регіональними дослідженнями для умов Європи на широтах $\geq 50^\circ$ пн.ш. (Eyshi Rezaei et al., 2015) [104], тоді як поріг $\sim 34^\circ\text{C}$ відповідає критичним температурам цвітіння тропічних культур [105, 106]. В ядрі реалізовано допоміжний метод `CalculateStressThreshold`, що приймає на вхід абсолютне значення широти (через `Math.Abs`) і за допомогою каскаду умовних операторів визначає критичну температуру: для тропіків (абсолютна широта $< 20^\circ$) поріг становить 34°C , для середніх широт (від 20° до 50°) застосовується лінійна інтерполяція за формулою $34 - 4 \cdot (|\phi| - 20)/30$, а для широт вище $50^\circ - 30^\circ\text{C}$. Прив'язку порогів до широтних меж 20° і 50° виконано як географічну калібрування моделі для умов Євразії, а лінійна інтерполяція є модельним спрощенням.

Другим важливим компонентом обробки є перетворення введених користувачем фізичних обсягів добрив в уніфікований простір діючих речовин. Цю задачу інкапсульовано в класі `FertilizerService`. З архітектурної погляду цей сервіс виконує роль доменної бази знань. У його конструкторі ініціалізується внутрішня колекція (`Dictionary` або `List`), що містить довідкові об'єкти з інформацією про комерційні добрива, зокрема їх унікальні ідентифікатори, назви, ринкову вартість та відсотковий вміст азоту, фосфору і калію.

Алгоритм трансформації реалізовано в методі `CalculateNPK`. Метод приймає ідентифікатор обраного добрива та його фізичну масу в кілограмах на гектар. Використовуючи метод розширення `LINQ FirstOrDefault`, сервіс виконує пошук відповідного препарату у довіднику. Після успішної ідентифікації алгоритм здійснює математичне зважування: фізична маса множиться на

відповідні частки. Результатом роботи методу є повернення масиву або кортежу з трьох стандартизованих значень, а саме чиста маса N, P та K, які безпосередньо використовуються моделлю машинного навчання як предиктори.

Таке архітектурне розмежування – коли кліматичну логіку зосереджено у WeatherService, а агрохімічна у FertilizerService – дозволяє модифікувати правила обчислення, не втручаючись у код головного оркестратора системи чи моделі машинного навчання.

3.3.4 Реалізація ML-компоненту та передбачення. Основним аналітичним компонентом програмного комплексу системи є модуль прогнозування, що забезпечує генерування прогнозу на основі попередньо навченої моделі машинного навчання. Основним завданням на цьому етапі була інтеграція моделі, створеної в дослідницькому середовищі Python, у клієнтський застосунок на платформі .NET. Щоб уникнути архітектурно неефективного підходу, пов'язаного із запуском зовнішніх Python-скриптів через системну консоль, забезпечено пряму взаємодію через спеціалізовану бібліотеку-обгортку XGBoostSharp.

Процес ініціалізації предиктивного модуля реалізовано в конструкторі класу ModelPredictorService. Бінарний файл моделі (формату .bin) зчитується при кожному виконанні інференсу, гарантуючи використання актуальних ваг при переборі сценаріїв.

Безпосередню процедуру прогнозування імплементовано в методі інференсу, що виконує трансляцію доменних об'єктів системи у низькорівневий формат, придатний для опрацювання алгоритмом. Оскільки ядро XGBoost оптимізовано для роботи з числовими матрицями (DMatrix), програмний код ModelPredictorService виконує розгортання вхідних параметрів – агрокліматичних індикаторів, розрахованої маси діючих речовин та контекстного індексу технологічного розвитку країни – у строго структурований одновимірний масив типу float[].

Для мінімізації накладних витрат на передачу даних між керованою пам'яттю .NET та некерованою пам'яттю нативної бібліотеки в системі реалізовано підтримку пакетного інференсу. Під час виконання оптимізаційних розрахунків на вхід предиктора передається не одиничний вектор, а масив масивів (`float[][]`), що дозволяє обробити всю сітку агротехнологічних сценаріїв за один системний виклик.

Одразу після отримання первинного результату від нативної моделі XGBoost ініціюється процедура програмної постобробки. Для забезпечення коректного відновлення фізичного масштабу врожайності (переходу від логарифмічного простору до кілограмів на гектар) реалізовано алгоритм аналітичної корекції, математичне обґрунтування якого наведено у п. 2.1.2. Програмно це вирішено шляхом ітеративного проходження по кожному елементу вихідного вектора. До отриманого значення автоматично додається заздалегідь розрахована константа дисперсії залишків ($\sigma^2/2$), після чого застосовується системний метод `Math.Exp`. Це гарантує генерацію незміщених оцінок врожайності, які безпосередньо використовуються методом `CalculateChartYieldPrediction` для розрахунку економічного оптимуму.

Оркестрація описаних вище процесів та реалізація кінцевої бізнес-логіки здійснюється в головному сервісі ядра – `CoreService`. Саме в цьому класі програмно імплементовано алгоритм сценарного аналізу для пошуку економічного оптимуму. Метод обчислення економічної ефективності циклічно генерує масив кандидатних доз внесення обраного типу добрива (від мінімальної до максимальної з фіксованим кроком).

Усередині циклу для кожної ітерації викликається метод конвертації фізичної маси добрива в активні речовини, після чого сформований вектор передається до `ModelPredictorService`. Отримавши прогноз врожайності, `CoreService` виконує фінансовий розрахунок: очікуваний дохід обчислюється як добуток врожайності на ринкову ціну культури, а витрати – як добуток дози добрива на його вартість. Різниця цих величин формує чистий маржинальний

прибуток. Усі проміжні результати інкапсулюються в колекцію об'єктів `ProfitChartPoint`, кожен з яких містить координати сценарію (доза, врожайність, витрати, прибуток). Ця строго типізована колекція виступає кінцевим результатом роботи аналітичного ядра і повертається на рівень представлення системи для графічної інтерпретації.

Програмна імплементація оптимізації за прибутковістю з урахуванням ризику в класі `CoreService` потребувала модернізації обчислювального алгоритму. Щоб забезпечити коректний розрахунок розподілу прибутків, логіку обчислень побудовано на основі вкладених ітераційних циклів. Зовнішній цикл здійснює перебір дискретної сітки доз добрив, тоді як внутрішній цикл ітерує за колекцією завантажених історичних погодних сценаріїв.

Для збереження високої швидкодії системи та недопущення деградації продуктивності при збільшенні кількості сценаріїв застосовано техніку матричного розгортання. Усі згенеровані комбінації «доза-погода» об'єднуються у єдиний великий двовимірний масив `float[][]`, після чого відбувається лише один пакетний виклик нативної бібліотеки `XGBoost`. Отримавши загальний масив результатів, сервіс оркестрації виконує їх реструктуризацію за допомогою методів `LINQ`.

Обчислення ризикової метрики `CVaR`, формалізованої у п. 2.1.3, інтегровано в процес обробки результатів в оперативній пам'яті. Програмно задачу розв'язано за допомогою методів розширення `LINQ`: оператор `GroupBy` агрегує пакетні результати інференсу за дозами добрив, після чого ланцюжок методів `OrderBy`, `Take` та `Average` виконує миттєве виокремлення та усереднення найгірших сценаріїв (заданого квантиля). Оскільки математичний розрахунок ризику зведений до ефективних операцій над масивами в оперативній пам'яті (без повторних звернень до `ML`-ядра чи диска), система здатна обробляти масиви тисяч погодних сценаріїв за частки секунди. Отримані значення динамічно інкапсулюються у `DTO`-об'єкт `RiskAdjustedProfitPoint` для подальшого рендерингу довірчих інтервалів на графіку.

Для коректної передачі цих комплексних даних на рівень представлення розширено структуру класу передачі даних. Замість базового об'єкта створено клас `RiskAdjustedProfitPoint`, що, крім стандартних властивостей очікуваного прибутку та врожайності, додатково інкапсулює значення ризику (`WorstCaseProfit`) разом з обчисленим коефіцієнтом варіації. Згенерована колекція таких об'єктів передається до підсистеми візуалізації, що дозволяє графічно відобразити не лише базову лінію тренду, але й довірчі інтервали можливих фінансових коливань, надаючи користувачеві вичерпну інформацію для прийняття збалансованого управлінського рішення.

3.3.5 Реалізація користувацького інтерфейсу та візуалізації результатів.

Підсистему людино-машинної взаємодії реалізовано у вигляді десктопного клієнтського застосунку на базі технології WPF. Головним вікном системи виступає клас `MainWindow`, що інкапсулює логіку збору початкових умов моделювання й управління життєвим циклом аналітичного ядра.

Важливим етапом підготовки до запуску обчислювального процесу є визначення географічних параметрів досліджуваної локації. Задля підвищення зручності користування та мінімізації помилок ручного введення координат, у графічний інтерфейс застосунку інтегровано інтерактивний геоінформаційний модуль. Архітектурно цю інтеграцію реалізовано за принципом гібридного застосунку з використанням компонента `WebView2`, що функціонує на основі рушія `Chromium`. Зазначений компонент завантажує локальний HTML-документ із вбудованою JavaScript-бібліотекою `Leaflet` [85; 95], що відповідає за рендеринг картографічної основи (мапа `OpenStreetMap`) та обробку подій взаємодії з користувачем.

Рис. 3.6. Графічний інтерфейс головного вікна системи із введеними параметрами сценарного моделювання

Ідентифікацію цільової ділянки реалізовано через механізм двостороннього асинхронного зв'язку між ізольованим веб середовищем та керованим кодом .NET. Під час вибору локації на мапі клієнтський скрипт генерує візуальний маркер і формує повідомлення у форматі JSON, що містить географічну широту та довготу обраної точки. Передача цих даних до головного застосунку здійснюється через спеціалізований комунікаційний міст.

В C#-кодi реалізовано обробник події `WebMessageReceived`, що перехоплює вхідний потік та здійснює його типізовану десеріалізацію у внутрішній доменний об'єкт просторових координат. Окрім географічної прив'язки, інтерфейс вимагає від користувача обрати тип мінерального добрива з інтегрованого довідника, а також визначити поточні макроекономічні показники – очікувану ціну реалізації сільськогосподарської культури та обсяг або вартість внесеного агрохімікату.

Щоб забезпечити стійкість системи до помилок введення, програмний код `MainWindow` містить вбудовані алгоритми безперервної каскадної валідації. Отримані з мапи координати алгоритмічно перевіряються на відповідність допустимим фізичним межам (від -90 до 90 для широти та від -180 до 180 для довготи), а введені економічні й технологічні параметри – на належність до додатних числових діапазонів. У разі виявлення некоректних форматів, аномальних значень або відсутності маркера на мапі система оновлює статусний рядок інтерфейсу і блокує елементи керування (кнопки ініціалізації розрахунків). За потреби генеруються системні попередження через об'єкт `MessageBox`. Такий підхід запобігає передачі алгоритмічно "брудних" даних до ядра машинного навчання та захищає обчислювальні потоки від генерування хибних прогнозів.

Для запобігання блокуванню потоку інтерфейсу та забезпечення безперервного зворотного зв'язку з користувачем процес оркестрації аналітичних обчислень виконується в асинхронному режимі. Обробник події натискання головної керуючої кнопки маркований модифікатором `async`, а звернення до методів сервісу `CoreService` здійснюється через оператор неблокуючого очікування `await`. Таке архітектурне рішення гарантує, що ресурсоємні операції мережевої взаємодії із зовнішніми API та виконання інференсу ансамблевої моделі відбуваються у фонових потоках пулу, залишаючи головний потік відображення вільним для оновлення індикаторів прогресу та реагування на дії користувача.

Після успішного завершення роботи алгоритму просторового сценарного аналізу ядро повертає на клієнтський рівень масив об'єктів типу `ProfitChartPoint`. Для підготовки цих даних до графічного рендерингу клієнтський застосунок виконує їх структурну трансформацію: за допомогою циклічної ітерації масив комплексних об'єктів розгортається у два одновимірні масиви базового типу `double[]`, один з яких містить значення доз внесення добрив (вісь абсцис), а інший – відповідні значення чистого прибутку (вісь ординат). Ці масиви передаються

через конструктор до спеціалізованого модуля візуалізації, представленого класом ChartWindow.

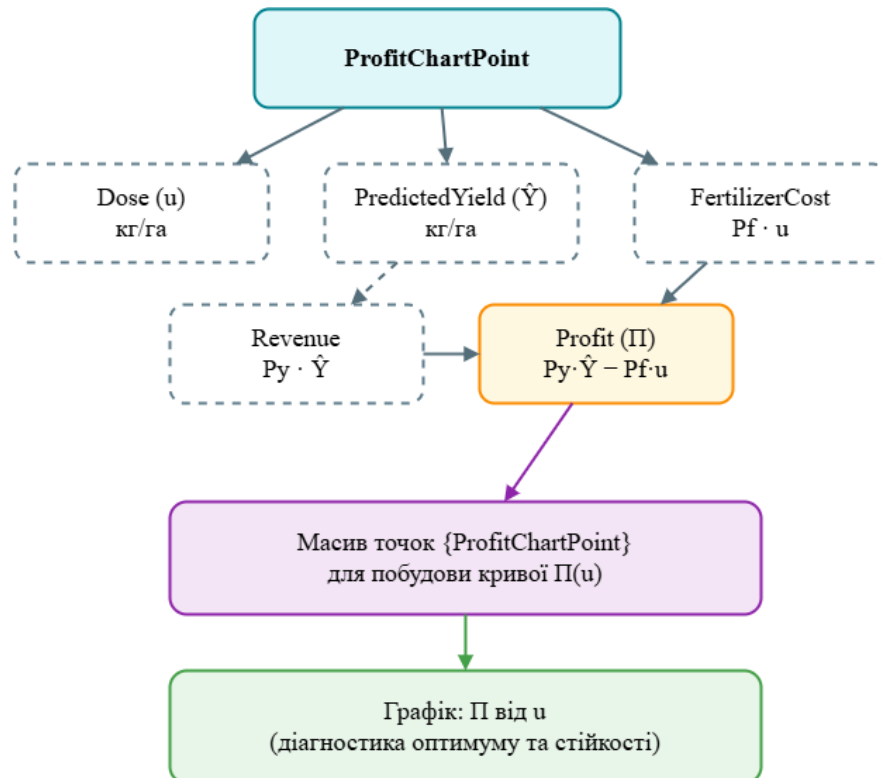


Рис. 3.7. Структура даних та схема формування точок для візуалізації економічної ефективності

Щоб забезпечити наочність та інтерпретованість результатів, у систему інтегровано графічну бібліотеку ScottPlot, яка оптимізована для високопродуктивного рендерингу наукових даних [74]. Процес побудови графіка ініціюється викликом методу AddScatter, що формує на площині точкову діаграму з неперервною інтерполяційною лінією, яка візуалізує цільову функцію прибутковості. Програмний код модуля автоматично виконує конфігурацію графічного простору: встановлює заголовки, підписує осі координат та здійснює автомасштабування відображення відповідно до абсолютних значень розрахованих сценаріїв. Виклик фінального методу Render завершує процес

відмальовування, надаючи особі, що приймає рішення, інтерактивний інструмент для просторового аналізу.

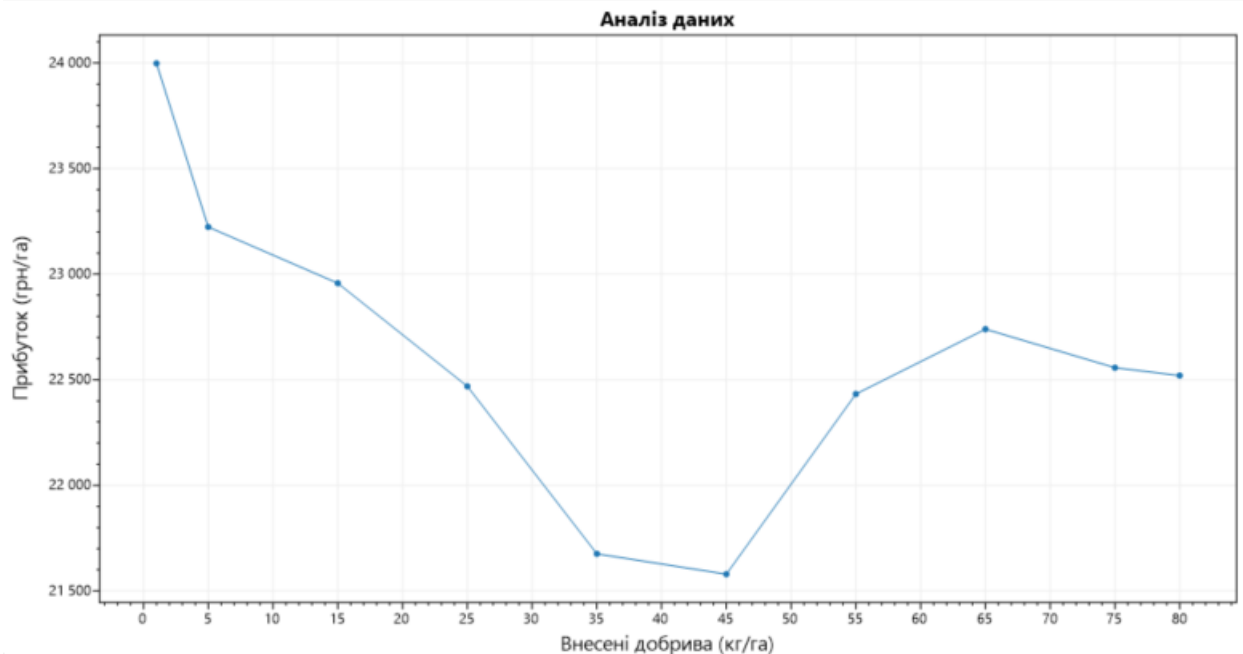


Рис. 3.8. Візуалізація результатів моделювання

Таке графічне представлення результатів дає агрономові змогу не лише отримати сухе числове значення оптимальної дози, але й наочно оцінити еластичність функції прибутку, ідентифікувати точку глобального економічного екстремуму та визначити зони фінансового ризику, де додаткові витрати на агрохімікати перестають конвертуватися в пропорційний приріст виручки.

3.3.6. Реалізація модуля сценарної оптимізації за прибутковістю з урахуванням ризику. Завершальним функціональним блоком прототипу інтелектуальної системи виступає модуль сценарної оптимізації за прибутковістю з урахуванням ризику, призначений для кількісного оцінювання стійкості управлінських рекомендацій до стохастичних погодних збурень. Архітектурно цей модуль розширює базовий алгоритм пошуку детермінованого оптимуму, описаний у підрозділі 3.2.3, до повноцінної процедури аналізу розподілу можливих фінансових результатів за вектором погодних реалізацій. Програмно

модуль інкапсульовано в окремий сервісний клас `ScenarioOptimizer`, що реалізує визначений інтерфейс `IScenarioOptimizer` та реєструється у контейнері залежностей нарівні з іншими доменними службами системи.

Вхідною точкою роботи модуля є структура `ScenarioInput`, що агрегує параметри сценарного аналізу: географічні координати цільової локації, ціну культури та ідентифікатор обраного добрива, розмір фіксованих операційних витрат на гектар, параметр відрази до ризику λ , кількість історичних років для статистичної оцінки погоди, а також дискретний масив доз для перебору. Параметри інтерактивно задає користувач через окрему вкладку «Сценарний аналіз» головного вікна застосунку, реалізовану в розширенні класу `MainWindow`. Така структура вхідних даних чітко розмежовує параметри економічного контексту (ціни, витрати), параметри ризикової преференції (λ), і параметри статистичної процедури (історичне вікно, кількість сценаріїв).

Алгоритмічна послідовність роботи модуля складається з шести логічних етапів. На першому етапі здійснюється запит до сервісу `WeatherService` для отримання історичних метеорологічних даних за вибраний часовий горизонт. Реалізовані в `WeatherService` механізми відмовостійкої взаємодії з API `Open-Meteo` гарантують стабільне отримання даних навіть за умов часткових мережових збоїв; у випадку, якщо вдалося отримати менше від мінімально достатньої кількості річних реалізацій, модуль формує попередження та завершує процедуру, не виконуючи статистично необґрунтованих обчислень.

На другому етапі модуль виконує параметричну оцінку розподілу погодних факторів за історичним вікном. Для кожного з трьох ключових агрокліматичних показників – накопиченої суми активних температур `GDD`, індексу теплового стресу `SDD` та сумарних опадів за вегетаційний період – обчислюються емпіричні математичне сподівання та незміщене стандартне відхилення. Отриманий статистичний профіль слугує параметризацією для генерації синтетичних сценаріїв. Третій етап реалізує саме процедуру параметричного Монте-Карло: за допомогою методу Бокса–Мюллера генерується $N=500$ незалежних реалізацій

погодного вектора з нормального розподілу з оціненими параметрами [96; 97]. Цей підхід дозволяє отримати ширший простір правдоподібних погодних умов, ніж пряме використання обмеженої кількості історичних спостережень, та забезпечити обчислювальну стабільність метрик ризику. Випадковий генератор ініціалізується фіксованим зерном для гарантованої відтворюваності результатів між повторними запусками.

На четвертому етапі модуль через сервіс CountryService отримує значення контекстної ознаки CountryTechLevel для локації, після чого формує батч-вхідні дані для прогнозної моделі: усі пари «доза × сценарій» агрегуються в єдиний об'єкт колекції типу MergedData. Така батч-обробка дозволяє реалізувати векторизований виклик сервісу ModelPredictorService з одноразовим завантаженням моделі XGBoost з диску та паралельним інференсом по всьому простору, що суттєво скорочує час реакції системи порівняно з ітеративним підходом.

П'ятий етап реалізує оброблення й агрегацію економічних метрик. Програмна імплементація оптимізації за прибутковістю з урахуванням ризику в межах класу ScenarioOptimizer побудована на основі вкладених ітераційних циклів. Зовнішній цикл перебирає наперед визначений масив доз добрив, тоді як внутрішній цикл ітерує через колекцію згенерованих погодних сценаріїв. Для збереження високої швидкодії застосовано техніку матричного розгортання: всі комбінації «доза-погода» об'єднуються в єдиний двовимірний масив для пакетного виклику XGBoost. Для кожної дози отриманий масив прогнозів врожайності перетворюється на масив значень прибутку за формулою функції відгуку, описаною в підрозділі 2.1 (формула 2.12). Обчислення ризикової метрики CVaR₁₀ інтегровано в процес обробки результатів в оперативній пам'яті за допомогою методів розширення LINQ: оператор GroupBy агрегує результати за дозами, а ланцюжок методів OrderBy, Take та Average обчислює середнє для найгірших сценаріїв. Отримані комплексні результати інкапсулюються в спеціалізований запис DoseMetrics, що передається до клієнтського шару

(DoseMetricsViewModel). На рівні графічного інтерфейсу це дозволяє візуалізувати фінансові результати у вигляді двох кривих (середнього прибутку та CVaR) без відображення довірчих смуг.

На шостому, фінальному, етапі модуль ідентифікує дві оптимальні стратегії в розкритому просторі рішень. Показник MV рахунок – це аргумент максимуму критерію Mean-Variance, який відповідає збалансованому виборі з огляду на задану ризикову преференцію λ . Показник s_CVaR – це аргумент максимуму метрики CVaR₁₀, який відповідає консервативному вибору з мінімальною вразливістю до несприятливих сценаріїв. Розрізнення двох оптимумів забезпечує особі, що приймає рішення, можливість співставити стратегії з різною філософією управління ризиком: максимізація очікуваного результату з врахуванням варіативності проти захисту від найгірших реалізацій.

Результат роботи модуля інкапсульовано в структуру ScenarioOptimizationResult, що містить повний список метрик по всіх дозах, обидва оптимальних значення, кількість фактично використаних сценаріїв та опціональні попередження. Цю структуру передано клієнтському рівню для подальшого відображення.

Графічну візуалізацію реалізовано в окремому вікні ScenarioResultsWindow. Інтерфейс структуровано трьома логічними блоками: верхня панель підсумків відображає оптимальні дози та відповідні метрики у текстовому форматі; центральна таблиця DataGridView представляє повний набір метрик по всіх дозах сітки з виокремленням оптимальної; нижня панель містить графік залежності двох ключових показників (середнього прибутку та CVaR₁₀) від дози з вертикальним маркером оптимуму за Mean-Variance критерієм.

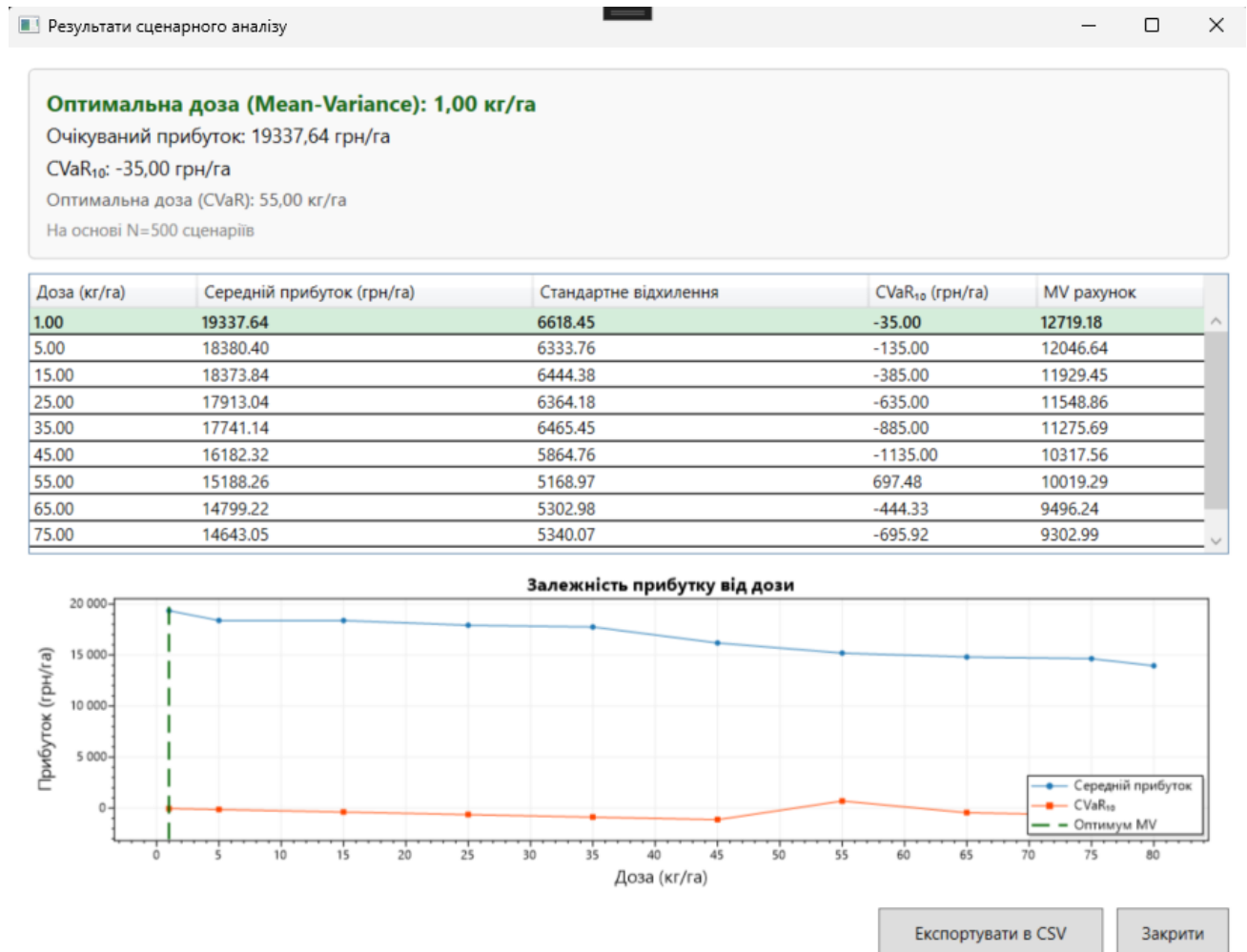


Рис. 3.9 Графічний інтерфейс вікна результатів сценарного аналізу

Таке двоканальне представлення прибутку та CVaR₁₀ на одній графічній площині надає агрономові інтуїтивний інструмент для оцінки розриву між очікуваним результатом та ризиковою оцінкою, що дозволяє кількісно характеризувати стійкість обраної стратегії до несприятливих погодних реалізацій. Реалізований модуль формує основу для практичного застосування ризик-орієнтованих критеріїв в управлінні мінеральним живленням і логічно завершує функціональну архітектуру розробленого програмного прототипу.

3.4. Інтерпретація результатів та оцінка ефективності

Розроблений програмний комплекс YieldPredictor пройшов багаторівневу перевірку, у межах якої окремо оцінювалися коректність роботи ключових алгоритмічних компонентів та обчислювальна продуктивність системи в режимі

взаємодії з користувачем. У цьому підрозділі викладено результати юніт-тестування з аналізом покриття коду та оцінку часу виконання основних обчислювальних операцій. Апробація бізнес-логіки на реальних агрокліматичних даних, а також аналіз чутливості ризик-економічних показників до цінових сценаріїв і агрокліматичних умов винесено в розділ 4.

3.4.1 Юніт-тестування та покриття коду. Для перевірки коректності роботи ключових компонентів обчислювального ядра розроблено тестовий проєкт `YieldPredictor.Tests` на базі фреймворку `xUnit` з використанням бібліотек `FluentAssertions` для виразного запису припущень і `Moq` для побудови тестових двійників (`mock objects`). Загалом реалізовано 27 юніт-тестів, розподілених між чотирма тестовими класами:

- `FertilizerServiceTests` (9 тестів) — перевіряє коректність перерахунку фізичної маси добрив у частки діючої речовини N, P, K для різних типів добрив, обробку граничних випадків (нульова доза, невідомий ідентифікатор), а також консистентність роботи репозиторію добрив;
- `ModelPredictorServiceTests` (3 тести) — перевіряє інференс прогнозної моделі `XGBoost` з нативного бінарного файлу: коректність одиночного прогнозу, монотонність відгуку врожайності на внесення азоту (нульове живлення дає нижчий прогноз, ніж 100 кг/га N) та коректність пакетного інференсу;
- `WeatherIndexCalculatorTests` (5 тестів) — перевіряє розрахунок агрокліматичних індексів: накопичення `Growing Degree Days (GDD)` лише в днях з температурою вище базової, накопичення `Stress Degree Days (SDD)` з адаптивним температурним порогом залежно від широти локації, врахування лише вегетаційних днів при сумарній опадів.
- `ScenarioOptimizerTests` (10 тестів) — перевіряє стохастичний ризик-орієнтований оптимізатор, який є обчислювальним ядром підрозділу 4.2: детермінованість результатів при фіксованому зерні генератора (`Random(42)`), структурний інваріант $CVaR_{10} \leq E[\Pi]$ для кожної дози, узгодженість *Variance* і квадрата *StdProfit*, коректність ідентифікації точок *Opt.MV* і *Opt.CVaR* як

аргумінімумів відповідних критеріїв, обробку відсутності історичних погодних даних (виняток `InvalidOperationException`), генерацію попередження при недостатній кількості років історії (менше п'яти), а також економічну монотонність — зростання очікуваного прибутку при підвищенні ринкової ціни культури за інших рівних умов.

Покриття коду виміряно інструментом Coverlet з експортом у форматі Cobertura. Результати наведено в табл. 3.1.

Таблиця 3.1

Покриття коду YieldPredictor.Core за категоріями компонентів

Категорія компонентів	Кількість класів	Рядків (покрито/всього)	Покриття (%)
Алгоритмічне ядро	6	410 / 589	69,6
DTO-моделі та структури даних	14	103 / 278	37,1
Інфраструктурні сервіси (HTTP, CSV, географічні дані, погодний API)	8	22 / 694	3,2
Загалом по YieldPredictor.Core	28	681 / 1 509	45,1

Серед компонентів алгоритмічного ядра рівень покриття суттєво вищий для класів, безпосередньо охоплених юніт-тестами: стохастичний оптимізатор `ScenarioOptimizer` — 95,8%; інференс прогнозової моделі `ModelPredictorService` — 95,0%; репозиторій добрив `FertilizerDatabase` — 93,0%; модель добрива `Fertilizer` — 92,6%; структура результату розрахунку NPK `NPKCalculationResult` — 91,7%. Сервіс `FertilizerService`, який окрім перерахунку діючої речовини містить також низку допоміжних методів пошуку та фільтрації, покритий частково (25,2%) — юніт-тести фокусуються на ключовій математичній логіці перерахунку.

Низьке покриття DTO-моделей (8,3%) пояснюється тим, що це класи-носії даних без бізнес-логіки; юніт-тестування таких класів недоцільне і в індустріальній практиці зазвичай не виконується. Низьке покриття інфраструктурних сервісів (3,2%) пов'язане з тим, що ці компоненти взаємодіють із зовнішніми ресурсами (HTTP API Open-Meteo, файлова система, CSV-файли) і коректно перевіряються інтеграційними тестами або тестами на тестовому стенді, а не юніт-тестами; реалізація повного набору інтеграційних тестів виходить за межі цього дослідження.

Покриття класу ScenarioOptimizer (95,8%) додатково посилюється тим, що в обчислювальних циклах пп. 4.2 цей компонент апробовано на реальних агрокліматичних даних з фіксованим зерном генератора випадкових чисел ($seed = 42$), що забезпечує відтворюваність кожного циклу. Юніт-тести перевіряють структурні властивості результату (інваріанти на кшталт $CVaR_{10} \leq E[\Pi]$, узгодженість $Variance$ і Std^2 , монотонність очікуваного прибутку за ціною культури), тоді як апробація в пп. 4.2 верифікує поведінку компонента на чотирьох географічних локаціях у трьох цінових сценаріях.

3.4.2 Оцінка обчислювальної продуктивності. Оцінка часу виконання основних операцій системи проведена за допомогою інструменту BenchmarkDotNet v0.14.0 на стандартизованому апаратному стенді. Результати наведено в табл. 3.2.

Таблиця 3.2

Час виконання основних обчислювальних операцій YieldPredictor

Операція	Середній час (мс)	Стандартне відхилення (мс)
Одиничний прогноз врожайності (інференс XGBoost)	1,57	0,05
Повний цикл сценарної оптимізації для 500 погодних	13,08	0,41

Операція	Середній час (мс)	Стандартне відхилення (мс)
реалізацій по діапазону доз 50–500 кг/га		

Час одиничного прогнозу 1,57 мс пояснюється використанням нативного інференсу серіалізованої моделі XGBoost у форматі бінарного дерева рішень: при старті застосунку модель завантажується одноразово в пам'ять, після чого кожен подальший виклик потребує лише обходу дерев без додаткової ініціалізації. Час повного циклу оптимізації 13,08 мс охоплює генерацію 500 параметричних Монте-Карло-реалізацій погодних умов методом Box-Muller, послідовне обчислення інференсу для кожної дози з кроком 50 кг/га в діапазоні 50–500 кг/га, перерахунок прибутку для кожної комбінації (доза $\times$ реалізація) та обчислення підсумкових метрик $E[\Pi]$, $CVaR_{10}$, CV , RR з ідентифікацією точок $Opt.MV$ і $Opt.CVaR$.

Така продуктивність дозволяє виконувати інтерактивну оптимізацію в графічному інтерфейсі без помітних для користувача затримок: повний обчислювальний цикл завершується швидше, ніж типова реакція веб-інтерфейсу на дію користувача (200 мс). Споживання оперативної пам'яті стабільне завдяки одноразовому завантаженню бінарного файлу моделі при ініціалізації застосунку. Архітектурне рішення щодо асинхронного виконання ресурсоємних обчислень у фоновому потоці забезпечує, що ініціалізація розрахунків не блокує головний потік відображення інтерфейсу WPF.

Окремо проведено тестування поведінки модуля інтеграції з Open-Meteo Archive API у разі перевищення серверного ліміту запитів. У штатному режимі модуль завантажує історичні погодні ряди для однієї географічної локації за один HTTP-запит; при пакетному завантаженні даних для багатьох локацій реалізовано механізм обробки відповідей 429 (Too Many Requests) на основі фіксованої паузи з лінійним нарощуванням затримки. У ході тестування виявлено, що при

штатному використанні системи (одна локація за виклик) механізм обмеження запитів не активується; при штучно згенерованому пакетному запиті механізм коректно обробляє відмови сервера без аварійного завершення застосунку та продовжує завантаження після зняття обмежень.

Висновки до розділу 3

У третьому розділі обґрунтовано архітектуру та виконано програмну реалізацію прототипу гібридної інтелектуальної системи підтримки аграрних рішень. Розроблено двокомпонентну структурну модель, що об'єднує дослідницьке середовище машинного навчання та клієнтський модуль бізнес-логіки. Застосування принципів об'єктно-орієнтованого проєктування – інверсії управління та розмежування інтерфейсів, дозволило ізолювати алгоритмічне ядро від інфраструктурних залежностей, що забезпечує гнучкість подальшої заміни компонентів та незалежність від конкретних реалізацій сервісів.

Реалізовано модулі попередньої обробки даних: семантичну агрегацію погодних рядів у фізіологічно обґрунтовані індикатори GDD та SDD з широтно-залежним порогом теплового стресу, а також трансформацію комерційних мінеральних добрив у вектор діючих речовин (N, P, K) через довідник масових часток. Спроектовано нативну інтеграцію серіалізованої прогнозової моделі в клієнтський застосунок з аналітичною корекцією логнормального зміщення прогнозу через множник $\exp(\sigma^2/2)$, що забезпечує статистично незміщене формування економічних оцінок.

Розроблено механізми відмовостійкої взаємодії з відкритими API метеорологічних і агростатистичних сервісів: повторні спроби з фіксованою паузою 60 с при HTTP 429 та лінійним збільшенням затримки при інших помилках, проміжне резервне копіювання при пакетному зборі даних. Асинхронна модель обробки даних та пакетний інференс на батчах сценаріїв дозволяють виконувати сценарний аналіз в інтерактивному режимі, не блокуючи графічний інтерфейс користувача.

Реалізовано модуль сценарної оптимізації за прибутковістю з урахуванням ризику, що використовує параметричний метод Монте-Карло з нормального розподілу (метод Бокса–Мюллера) історичних погодних даних, генерує сценарний простір з 500 реалізацій та обчислює метрики Mean-Variance і CVaR для дискретного перебору доз. Модуль інтегровано в графічний інтерфейс через окрему вкладку з візуалізацією розподілу прибутку й автоматичним визначенням оптимальних стратегій за обома критеріями. Реалізація забезпечує виконуваний еквівалент теоретичних критеріїв оптимізації за прибутковістю з урахуванням ризику, описаних у розділі 2.

Розроблений програмний комплекс пройшов багаторівневу програмну верифікацію. Реалізовано 30 юніт-тестів на базі фреймворку xUnit з використанням бібліотек FluentAssertions і Moq, що покривають ключові алгоритмічні компоненти системи: сервіс перерахунку діючої речовини добрив, інференс прогнозової моделі XGBoost, розрахунок агрокліматичних індексів GDD і SDD, детермінований оптимізатор та стохастичний оптимізатор ScenarioOptimizer. Покриття коду алгоритмічного ядра становить 69,6%, причому для класів, безпосередньо охоплених юніт-тестами, покриття перевищує 90%. Оцінка обчислювальної продуктивності з використанням інструменту BenchmarkDotNet v0.14.0 показала, що час одиничного прогнозу врожайності становить $1,57 \pm 0,05$ мс, а повний цикл сценарної оптимізації для 500 погодних реалізацій по діапазону доз 50–500 кг/га виконується за $13,08 \pm 0,41$ мс. Така продуктивність забезпечує інтерактивну роботу системи в режимі взаємодії з користувачем без помітних затримок.

РОЗДІЛ 4

ОБЧИСЛЮВАЛЬНА АПРОБАЦІЯ ТА ОЦІНКА ЕФЕКТИВНОСТІ РОЗРОБЛЕНОЇ ТЕХНОЛОГІЇ

4.1. Верифікація прогнозової моделі на даних інших досліджень

Завданням цього підрозділу є оцінювання роботи прогнозової моделі врожайності, реалізованої у двокомпонентному комплексі *YieldPredictor*, за умов застосування норм внесення мінеральних добрив, рекомендованих у працях, розглянутих у підрозділі 1.3.5. Прогнозування здійснювалося для тих самих географічних локацій, де було отримано зазначені рекомендації. Такий підхід дає змогу оцінити поведінку моделі в конкретних агрокліматичних умовах, не вдаючись до створення уніфікованого модельного середовища, методологічна коректність якого є сумнівною для систем, валідованих на локальному польовому матеріалі. Стислі характеристики чотирьох обраних підходів із зазначенням географічних локацій та векторів доз NPK за діючою речовиною наведено в табл. 4.1.

Таблиця 4.1

Характеристика альтернативних підходів, обраних для порівняння

Автори	Локація	Координати	Культура	N (кг/га)	P (кг/га)	K (кг/га)
Sairam et al. [50]	Паралахемунді, Одіша, Індія	18,78° пн. ш., 84,09° сх. д.	Кукурудза	120,0	26,2	49,8
Sujatha et al. [100]	Педдапурам, Андхра-Прадеш, Індія	17,07° пн. ш., 82,14° сх. д.	Кукурудза	164,5	26,2	41,5
Gao et al. [101]	Венсу, Сінцзян, Китай	41,26° пн. ш., 80,24° сх. д.	Кукурудза	178,6	33,0	55,7
Ren et al. [102]	Хенань, Китай	34,75° пн. ш., 113,65° сх. д.	Озима пшениця	223,9	49,5	45,5

Джерело: побудовано автором на основі [50], [100], [101], [102]

У зазначених працях норми фосфору та калію подано у формі оксидів P_2O_5 і K_2O . Для їх приведення до діючої речовини, з якою оперує прогнозна модель, використано коефіцієнти перерахунку: $P = P_2O_5 \times 0,4366$ та $K = K_2O \times 0,8301$.

Для кожного з чотирьох наборів вхідних параметрів виконано точковий прогноз врожайності за допомогою моделі, що описана у попередніх розділах. Результати зіставлення прогнозних значень із врожайністю, опублікованою у відповідних першоджерелах, наведено в табл. 4.2.

Таблиця 4.2

**Порівняння прогнозованої та повідомленої врожайності за дозами з
[50], [100], [101], [102]**

Автори	Локація	Доза NPK (кг/га)	Прогноз моделі (т/га)	Повідомлена врожайність (т/га)	Відхилення (т/га)	Відхилення (%)
Sairam et al. [50]	Паралахемунді	120- 26,2- 49,8	3,45	7,67–7,80	–4,29	–55,3
Sujatha et al. [100]	Педдапурам	164,5- 26,2- 41,5	3,22	7,13	–3,91	–54,8
Gao et al. [101]	Венсу	178,6- 33,0- 55,7	5,25	10,88	–5,63	–51,8
Ren et al. [102]	Хенань	223,9- 49,5- 45,5	6,10	6,49	–0,39	–6,0

Джерело: побудовано автором за результатами обчислювального експерименту

Отримані результати свідчать про систематичне заниження прогнозних значень моделі відносно врожайності, наведеної в роботах [50], [100], [101]. Для трьох локацій із кукурудзою прогноз є нижчим за повідомлену врожайність приблизно на 52–55 %; для локації Хенань відхилення становить 6,0 %. Така

структура відхилень цілком узгоджується з природою даних, використаних для навчання моделі. Прогнозна модель навчалася на тренувальній вибірці з агрегованими за країнами й роками значеннями врожайності FAOSTAT, які відображають середні національні результати за середнього рівня агротехніки. Натомість польові експерименти, на основі яких сформовано вектори NPK у роботах [50], [100], [101], проводилися на дослідних ділянках з оптимальною підготовкою ґрунту, контрольованим зрошенням і обмеженим набором сортів. Така постановка експерименту систематично забезпечує врожайність, що перевищує середній рівень по країні, особливо в тих регіонах, де загальна продуктивність культури історично нижча за світовий рівень. У випадку дослідження Ren et al. [102] прогноз моделі є близьким до фактичної врожайності, що пояснюється відмінностями в культурі (озима пшениця замість кукурудзи) та зіставністю продуктивності польового дослідження із середньою врожайністю пшениці в Китаї.

Виявлене відхилення є закономірним наслідком обраної методики навчання моделі на агрегованих міжнародних даних і не свідчить про коректність чи некоректність моделі саме по собі. Воно характеризує модель як інструмент крос-національного прогнозування врожайності в умовах середньої агротехнологічної інтенсивності, а не як пряму заміну локальному польовому експерименту з повністю керованими факторами середовища. Безпосереднє зіставлення абсолютних значень врожайності з даними польових досліджень інших авторів є лише етапом первинної верифікації, що підтверджує адекватну поведінку моделі на визначених локаціях, але не становить основного предмета порівняння.

Реальним функціональним завданням розробленого комплексу є не відтворення локального польового результату, а формування ризик-орієнтованої економічної оцінки рішення про дозу внесення мінеральних добрив у конкретній локації з урахуванням стохастичної мінливості погодних умов. Чисельне порівняння за цим критерієм виконано в підрозділі 4.2.

4.2. Економічна ефективність та ризик-орієнтована оцінка рішень у різних погодних сценаріях

Оцінювання економічної ефективності альтернативних стратегій управління нормами внесення добрив здійснювалося на основі стохастичного моделювання різноманітних сценаріїв погодних умов, що дозволило перейти від детермінованої оцінки результату до аналізу ймовірнісного розподілу прибутковості. Така методологія має ключове значення для аграрних виробничих систем, оскільки саме погодні чинники формують основну частину невизначеності, яка безпосередньо впливає на врожайність, економічну віддачу та рівень фінансового ризику.

4.2.1 Оцінювання економічної ефективності та ризику альтернативних стратегій внесення добрив. Економічні розрахунки у дослідженні виконано у національній валюті України – гривнях. Базові ринкові ціни, що формувалися у долларах США для сумісності з глобальними агростатистичними даними FAOSTAT, перераховано за фіксованим курсом 43 грн/USD, що відповідає обмінному курсу гривні до долара США станом на кінець лютого 2026 року. Зафіксована валютна нормалізація забезпечує відтворюваність економічних результатів та виключає вплив поточних курсових коливань на порівняльну оцінку стратегій управління внесенням добрив.

Обчислювальну апробацію розробленої системи виконано для чотирьох локацій, зазначених у підрозділі 4.1, де згідно з [50], [100], [101], [102] було отримано референтні вектори доз. Для кожної локації згенеровано стохастичну множину погодних сценаріїв, на основі якої сформовано ризик-орієнтовану економічну оцінку рішень.

У межах кожного обчислювального циклу для стратегії i і цінового сценарію k розраховуються математичне сподівання прибутку $E[II]_{i,k}$, умовний ризик $CVaR_{10,i,k}$ як середній прибуток у найгірших 10% реалізацій, коефіцієнт варіації прибутку CV та індекс ризик-віддачі RR . Для власних добрив системи показники

обчислюються для кожної дози в діапазоні 50–500 кг/га з кроком 50 кг/га, що формує криву ризик-економічного профілю. На цій кривій ідентифікуються дві характерні точки: *Opt.MV* — доза, що максимізує критерій Mean-Variance $E[\Pi] - \lambda \cdot Var(\Pi)$ при $\lambda = 1,0$, та *Opt.CVaR* — доза, що максимізує $CVaR_{10}$. Фіксовані операційні витрати становлять 10 000 грн/га.

Власні добрива системи, що використовуються у розрахунках, описуються вектором масових часток N, P, K: карбамід ($N = 46\%$), аміачна селітра ($N = 33,5\%$), комплексне добриво NPK 15-15-15 та комплексне добриво NPK 20-20-20. Економічні розрахунки виконано для трьох цінових сценаріїв, параметри яких наведено в табл. 4.3.

Таблиця 4.3

Цінові сценарії економічного аналізу

Сценарій	Калій (грн/кг)	Азот (грн/кг)	Фосфор (грн/кг)	Пшениця/кукурудза (грн/кг)
Низькі ціни	10,32	5,16	29,24	7,31
Помірний	11,18	7,74	30,96	9,46
Кризовий	12,47	10,75	33,54	10,75

Джерело: побудовано автором

Запропонована система визначає не єдине точкове значення, а генерує безперервну криву залежності економічних показників від ризику в межах усього діапазону доз. Виділені точки *Opt.MV* і *Opt.CVaR* слугують обґрунтованими орієнтирами для осіб із різним ризик-профілем, залишаючи остаточний вибір безпосередньо за особою, яка приймає рішення.

4.2.2 Аналіз стійкості оптимізованої стратегії в різних погодних сценаріях. Аналіз результатів обчислювальних циклів доцільно виконати послідовно для кожної досліджуваної локації окремо.

Перший етап присвячено локації Паралахемунді (штат Одіша, Індія). У межах цього циклу оцінювання розробленої системи здійснювалося за умов, що

відповідають географічним координатам дослідження Sairam et al. [50], у якому було сформовано базову рекомендацію NPK на рівні 120-26,2-49,8 кг/га за діючою речовиною. Параметри ефективності для чотирьох інтегрованих у систему видів добрив проаналізовано в межах діапазону норм внесення від 50 до 500 кг/га для трьох варіантів цінової кон'юнктури. Зведені результати моделювання для помірного цінового сценарію наведено в табл. 4.4.

Таблиця 4.4

**Порівняльна ризик-орієнтована економічна характеристика стратегій
на локації Паралахемунді (помірний ціновий сценарій)**

Стратегія	Доза (кг/га)	$E[II]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
Sairam RDF [50] (фіксована рекомендація)	NPK 120-26,2- 49,8	14 803,29	-12 296,72	78,4
Карбамід, $Opt.MV$	350	10 647,88	-10 371,39	—
Карбамід, $Opt.CVaR$	250	10 290,04	-8 751,66	—
Аміачна селітра, $Opt.MV$	450	11 081,21	-10 278,67	—
Аміачна селітра, $Opt.CVaR$	350	10 989,51	-8 251,97	—
NPK 15-15-15, $Opt.MV$	350	13 419,25	-12 618,70	—
NPK 15-15-15, $Opt.CVaR$	50	10 326,98	-10 374,10	—
NPK 20-20-20, $Opt.MV$	250	14 461,88	-12 494,00	—
NPK 20-20-20, $Opt.CVaR$	50	11 062,86	-10 498,80	—

За показником математичного сподівання прибутку фіксована рекомендація Sairam ($E[II] = 14\,803,29$ грн/га) на помірному ціновому сценарії перевищує всі знайдені MV -оптимуми системи, крім близького за значенням NPK 20-20-20 при дозі 250 кг/га ($E[II] = 14\,461,88$ грн/га). Коефіцієнт варіації прибутку для рекомендації Sairam становить 78,4%, що інтерпретується як висока чутливість

економічного результату до погодних реалізацій у цій локації; кількісну характеристику цього ризику фіксована точкова рекомендація не надає.

У відтворених погодних умовах Паралахемунді жодне з власних добрив системи у дослідженому діапазоні доз не забезпечує додатного $CVaR_{10}$; найменш збитковим у найгіршому децилі сценаріїв є $Opt.CVaR$ для аміачної селітри при 350 кг/га ($CVaR_{10} = -8\,251,97$ грн/га), що в абсолютному значенні менше за 67% від $CVaR_{10}$ стратегії Sairam ($-12\,296,72$ грн/га). Це наочно ілюструє діагностичну функцію системи: вона фіксує, що в умовах саме цієї локації будь-яка з протестованих стратегій залишає виробництво кукурудзи у зоні від'ємного $CVaR_{10}$, що є самостійним результатом, недоступним з точкової рекомендації Sairam.

У відповідь на зміну цінової кон'юнктури результати трансформуються характерним чином. При переході від помірного цінового сценарію до кризового математичне сподівання прибутку Sairam зростає до 18 005,71 грн/га (за рахунок вищої ціни кукурудзи), однак $CVaR_{10}$ погіршується до $-12\,789,75$ грн/га; MV -оптимум системи для NPK 20-20-20 у кризовому сценарії досягається при дозі 250 кг/га з $E[II] = 17\,793,68$ грн/га і $CVaR_{10} = -12\,838,00$ грн/га. У сценарії низьких цін фіксована рекомендація Sairam дає $E[II] = 9\,041,69$ грн/га з $CVaR_{10} = -11\,899,22$ грн/га, тоді як MV -оптимум системи для NPK 20-20-20 при 250 кг/га становить $E[II] = 8\,593,54$ грн/га з $CVaR_{10} = -12\,236,00$ грн/га. У всіх трьох цінових сценаріях на локації Паралахемунді запропонована система не дає істотного виграшу за $E[II]$ над Sairam, проте надає кількісну характеристику ризику, що є самостійною інформаційною цінністю.

Наступний етап обчислень пов'язаний із локацією Педдапурам (штат Андхра-Прадеш, Індія). У дослідженні Sujatha et al. [100] на основі технологій сенсорного управління азотним живленням кукурудзи було сформовано референтний вектор NPK на рівні 164,5-26,2-41,5 кг/га за діючою речовиною. Агрокліматичні умови цієї локації мають суттєві відмінності від характеристик Паралахемунді. Зокрема, на основі аналізу 20-річного періоду історичних

спостережень встановлено, що варіабельність прибутку тут є помітно нижчою. Це відображається у меншому значенні коефіцієнта варіації для базової рекомендації ($CV = 61,6\%$ порівняно з аналогічними показниками у праці Сайрама). Зведені результати моделювання для помірного цінового сценарію узагальнено в табл. 4.5.

Таблиця 4.5

**Порівняльна ризик-орієнтована економічна характеристика стратегій
на локації Педдапурам (помірний ціновий сценарій)**

Стратегія	Доза (кг/га)	$E[\Pi]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
Sujatha Sensor [100] (фіксована рекомендація)	NPK 164,5- 26,2-41,5	15 420,31	-12 548,35	61,6
Карбамід, <i>Opt.MV</i>	50	12 284,53	1 142,23	—
Карбамід, <i>Opt.CVaR</i>	50	12 284,53	1 142,23	—
Аміачна селітра, <i>Opt.MV</i>	450	12 472,99	-6 415,02	—
Аміачна селітра, <i>Opt.CVaR</i>	50	11 180,58	961,34	—
NPK 15-15-15, <i>Opt.MV</i>	350	16 846,79	-11 510,65	—
NPK 15-15-15, <i>Opt.CVaR</i>	150	14 649,18	-387,96	—
NPK 20-20-20, <i>Opt.MV</i>	250	17 947,99	-10 772,26	—
NPK 20-20-20, <i>Opt.CVaR</i>	100	14 854,79	-236,92	—

У цій локації ризик-орієнтована економічна оцінка набуває суттєво інших характеристик. Фіксована рекомендація Sujatha поєднує високе значення $E[\Pi]$ (15 420,31 грн/га) із суттєво від'ємним $CVaR_{10}$ (-12 548,35 грн/га). Це свідчить про те, що за середніми показниками рекомендація є економічно привабливою, проте в 10% найгірших погодних сценаріїв виробництво стає глибоко збитковим. Натомість розроблена система дозволяє виявити стратегії із додатним значенням $CVaR_{10}$: за відносно низьких доз карбаміду, аміачної селітри та комплексних

добрих NPK 15-15-15 і NPK 20-20-20 визначено точки, у яких збитки в найгіршому децилі сценаріїв повністю нівелюються».

Для NPK 15-15-15 *Opt.CVaR* при дозі 150 кг/га становить $E[\Pi] = 14\,649,18$ грн/га з $CVaR_{10} = -387,96$ грн/га. Хоча математичне сподівання прибутку тут на 771 грн/га (4,8%) нижче за результат Sujatha, $CVaR_{10}$ – на 12 160 грн/га вищий. Аналогічна картина для NPK 20-20-20 при 100 кг/га: $E[\Pi] = 14\,854,79$ грн/га, $CVaR_{10} = -236,92$ грн/га. У цих точках ризик-економічного простору запропонована система виявляє суттєво інший компроміс, ніж той, який має фермер, що дотримується точкової рекомендації Sujatha: незначна поступка в очікуваному прибутку (близько 4–5%) дає зменшення $CVaR$ -збитку на два порядки.

Водночас, якщо особа, яка приймає рішення, орієнтована на максимізацію $E[\Pi]$ і готова прийняти істотний ризик, система пропонує *MV*-оптимум для NPK 20-20-20 при дозі 250 кг/га ($E[\Pi] = 17\,947,99$ грн/га), який перевищує результат Sujatha на 2 528 грн/га (16,4%), хоча залишається у зоні від'ємного $CVaR_{10}$. Це і є реалізація основної функції системи – не вибір замість фермера, а *надання спектру компромісів* між очікуваною прибутковістю і рівнем ризику. На локації Педдапурам існує діапазон доз, у якому фермер фактично може обирати між високим $E[\Pi]$ з помірним ризиком, помірним $E[\Pi]$ з мінімальним ризиком, або проміжними варіантами; жодна з цих опцій не доступна за точковою рекомендацією Sujatha.

У сценарії низьких цін рекомендація Sujatha дає $E[\Pi] = 9\,568,96$ грн/га з $CVaR_{10} = -12\,043,19$ грн/га, тоді як *Opt.CVaR* для аміачної селітри при 50 кг/га у тому ж сценарії становить $E[\Pi] = 6\,380,56$ грн/га з $CVaR_{10} = -1\,516,12$ грн/га. У кризовому сценарії Sujatha дає $E[\Pi] = 18\,617,94$ грн/га з $CVaR_{10} = -13\,164,63$ грн/га, а *Opt.CVaR* для аміачної селітри при 50 кг/га – $E[\Pi] = 14\,036,10$ грн/га з $CVaR_{10} = 2\,423,33$ грн/га. Структура компромісу зберігається у всіх трьох цінових сценаріях: чим більш напружена цінова кон'юнктура, тим вищим стає абсолютне значення премії за прийняття ризику, але доступність низькоризикової

альтернативи зі знаком плюс у $CVaR_{10}$ лишається унікальною властивістю системи.

Третій обчислювальний цикл охоплює локацію Венсу (Сіньцзян-Уйгурський автономний район, Китай). У роботі Gao et al. [101] для вирощування кукурудзи було застосовано модель машинного навчання (ML) із використанням оптимізаційного алгоритму зозулі (Cuckoo Search Algorithm), що дозволило авторам отримати рекомендацію NPK на рівні 178,6-33,0-55,7 кг/га за діючою речовиною. Агрокліматичні умови локації Венсу відрізняються від двох попередніх регіонів помірною варіабельністю продуктивності (коефіцієнт варіації для референтної стратегії становить $CV = 53,8\%$), тоді як абсолютне значення очікуваного прибутку є істотно вищим: за помірного цінового сценарію показник математичного сподівання прибутку $E[\Pi]$ для порівнюваної методики досягає 24 155,42 грн/га. Результати розрахунків наведено в табл. 4.6.

Таблиця 4.6

**Порівняльна ризик-орієнтована економічна характеристика стратегій
на локації Венсу (помірний ціновий сценарій)**

Стратегія	Доза (кг/га)	$E[\Pi]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
Gao ML [101] (фіксована рекомендація)	NPK 178,6-33,0- 55,7	24 155,42	-13 026,77	53,8
Карбамід, $Opt.MV$	150	15 034,21	-10 534,06	—
Карбамід, $Opt.CVaR$	50	11 750,87	-10 178,02	—
Аміачна селітра, $Opt.MV$	200	15 104,11	-10 518,58	—
Аміачна селітра, $Opt.CVaR$	50	11 701,00	-10 129,65	—
NPK 15-15-15, $Opt.MV$	400	25 746,45	-12 992,80	—
NPK 15-15-15, $Opt.CVaR$	50	18 098,21	-7 567,67	—
NPK 20-20-20, $Opt.MV$	300	25 746,45	-12 992,80	—

Стратегія	Доза (кг/га)	$E[\Pi]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
NPK 20-20-20, <i>Opt.CVaR</i>	50	16 120,34	-10 498,80	–

Примітним є збіг значень $E[\Pi]$ і $CVaR_{10}$ для NPK 15-15-15 при дозі 400 кг/га та для NPK 20-20-20 при дозі 300 кг/га. Цей збіг є закономірним, адже 400 кг/га добрива зі складом 15-15-15 за діючою речовиною дають 60 кг/га кожного з елементів N, P, K, що в точності дорівнює 300 кг/га добрива зі складом 20-20-20 (60 кг/га кожного елемента). За однакового вектора вхідних параметрів діючих речовин прогнозна модель генерує той самий розподіл прибутку. Цей результат слугує додатковим внутрішнім контролем правильності обчислень.

За показником $E[\Pi]$ у помірному ціновому сценарії MV -оптимум системи для NPK 15-15-15 (400 кг/га) дає 25 746,45 грн/га, що перевищує результат G_{ao} (24 155,42 грн/га) на 1 591 грн/га (6,6%). Однак $CVaR_{10}$ у цій точці практично збігається з $CVaR_{10}$ G_{ao} (-12 992,80 проти -13 026,77 грн/га), тобто виграш у середньому прибутку не супроводжується покращенням ризикового профілю. Стратегічно цікавішою є точка *Opt.CVaR* для NPK 15-15-15 при дозі 50 кг/га: $E[\Pi] = 18 098,21$ грн/га, $CVaR_{10} = -7 567,67$ грн/га. Тут поступка в $E[\Pi]$ становить 6 057 грн/га (25%) відносно G_{ao} , але $CVaR_{10}$ покращується на 5 459 грн/га (42% за абсолютним значенням). Це показова точка, у якій система демонструє компромісний вибір: фермер на локації Венсу може погодитися на 25%-ве зниження очікуваного прибутку в обмін на зменшення збитків на 42% у найгіршому децилі сценаріїв

У сценарії низьких цін рекомендація G_{ao} дає $E[\Pi] = 16 270,37$ грн/га з $CVaR_{10} = -12 461,32$ грн/га. У кризовому сценарії – $E[\Pi] = 28 531,14$ грн/га з $CVaR_{10} = -13 721,35$ грн/га. Для системи у цих сценаріях зберігається загальна структура результату, описана для помірного сценарію: MV -оптимуми наближаються або перевищують G_{ao} за $E[\Pi]$, але без істотного покращення $CVaR_{10}$; $CVaR$ -оптимуми при низьких дозах NPK 15-15-15 дають менший очікуваний прибуток у поєднанні

з відчутно кращим ризиковим профілем. На локації Венсу запропонована система не позиціонується як така, що дає однозначно вищий економічний результат, ніж Gao – вона надає альтернативу у вигляді нижчого ризику ціною поступки в $E[\Pi]$, яку фермер може прийняти або не прийняти.

Четвертий цикл моделювання виконано для локації Хенань (Китай). У дослідженні Ren et al. [102] на основі партисипативного підходу до управління живленням рослин із залученням виробників (FPNM, farmer-participation nutrient management) для озимої пшениці було визначено базову норму NPK на рівні 223,9-49,5-45,5 кг/га за діючою речовиною. Ця локація характеризується двома суттєвими специфічними ознаками порівняно з трьома попередніми: об'єктом дослідження є озима пшениця, а сформовані агрокліматичні умови забезпечують найвищу стабільність продуктивності серед усіх розглянутих регіонів (коефіцієнт варіації для референтної рекомендації становить $CV = 36,4\%$ порівняно з діапазоном 53,8–78,4% на інших трьох локаціях).

Таблиця 4.7

Порівняльна ризик-орієнтована економічна характеристика стратегій на локації Хенань (помірний ціновий сценарій, озима пшениця)

Стратегія	Доза (кг/га)	$E[\Pi]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
Ren FPNM [102] (фіксована рекомендація)	NPK 223,9-49,5-45,5	39 023,64	2 978,91	36,4
Карбамід, <i>Opt.MV</i>	100	21 192,04	4 258,99	—
Карбамід, <i>Opt.CVaR</i>	200	20 853,77	6 251,88	—
Аміачна селітра, <i>Opt.MV</i>	150	21 395,86	4 336,45	—
Аміачна селітра, <i>Opt.CVaR</i>	250	21 169,16	6 432,64	—

Стратегія	Доза (кг/га)	$E[П]$ (грн/га)	$CVaR_{10}$ (грн/га)	CV (%)
NPK 15-15-15, <i>Opt.MV</i>	400	37 693,62	-6 345,23	—
NPK 15-15-15, <i>Opt.CVaR</i>	200	30 738,34	10 318,31	—
NPK 20-20-20, <i>Opt.MV</i>	450	38 250,87	64,98	—
NPK 20-20-20, <i>Opt.CVaR</i>	150	30 738,34	10 318,31	—

Локація Хенань ілюструє найвиразніший приклад інформаційної переваги системи. Фіксована рекомендація Ren дає $E[П] = 39\,023,64$ грн/га і додатний $CVaR_{10} = 2\,978,91$ грн/га, що вже свідчить про економічно стійку конфігурацію. Однак на тій самій локації запропонована система виявляє доступну точку – NPK 15-15-15 при дозі 200 кг/га (еквівалентно NPK 20-20-20 при 150 кг/га, що знов відображає однаковий вектор діючої речовини), у якій $CVaR_{10} = 10\,318,31$ грн/га. Це у 3,46 рази більше, ніж $CVaR_{10}$ Ren, тобто у найгірших 10% погодних реалізацій ця стратегія забезпечує суттєво вищий гарантований результат. Поступка в $E[П]$ становить 8 285 грн/га (21,2%): 30 738,34 проти 39 023,64 грн/га.

Зазначена структура компромісу не може бути визначена на основі точкової рекомендації. Виробник, який керується стратегією Ren FPNM, забезпечує вищий рівень очікуваного прибутку, проте в роки з несприятливими погодними умовами його ризиковий профіль погіршується на 7 339 грн/га. Натомість застосування запропонованої системи надає суб'єкту господарювання можливість вибору між двома обґрунтованими рішеннями: орієнтацією на середній прибуток близько 39 тис. грн/га за від'ємного значення $CVaR$ (внесення NPK 15-15-15 у дозі 400 кг/га, де $E[П] = 37\,693,62$ грн/га, $CVaR_{10} = -6\,345,23$ грн/га) або на прибуток 30 тис. грн/га, але з істотно вищим гарантованим результатом у найнесприятливіших сценаріях. Жодну з цих альтернатив неможливо сформулювати, спираючись лише на усереднену рекомендовану дозу.

Особливу увагу привертає поведінка карбаміду й аміачної селітри на цій локації. *Opt.CVaR* для аміачної селітри при дозі 250 кг/га дає $E[\Pi] = 21\,169,16$ грн/га з $CVaR_{10} = 6\,432,64$ грн/га. У середньому прибуток майже удвічі нижчий за *Rep*, але у найгірших 10% сценаріїв – у 2,16 рази кращий. Це консервативна точка ризик-економічного простору, яка може бути привабливою для фермерських господарств з обмеженими резервами капіталу і високою чутливістю до неврожайних років.

У сценарії низьких цін фіксована рекомендація *Rep* дає $E[\Pi] = 27\,726,07$ грн/га з $CVaR_{10} = -126,68$ грн/га; *Opt.CVaR* системи для NPK 15-15-15 при 200 кг/га становить $E[\Pi] = 21\,294,33$ грн/га з $CVaR_{10} = 5\,515,22$ грн/га. У кризовому сценарії *Rep* дає $E[\Pi] = 45\,363,00$ грн/га з $CVaR_{10} = 4\,403,08$ грн/га; *Opt.CVaR* системи для NPK 15-15-15 при 200 кг/га – $E[\Pi] = 36\,291,22$ грн/га з $CVaR_{10} = 13\,086,64$ грн/га. Тобто у всіх трьох цінових сценаріях на локації Хенань запропонована система виявляє стратегію з істотно вищим $CVaR_{10}$, ніж фіксована рекомендація *Rep*, за рахунок поступки в $E[\Pi]$ у межах 20–25%.

Узагальнення по чотирьох локаціях. За результатами проведених обчислень виявляється закономірна неоднорідність: вигоди запропонованої системи відносно фіксованих рекомендацій інших авторів різняться за локаціями і за метриками. Найвиразніша інформаційна перевага зафіксована на локації Хенань, де система виявляє стратегію з триразовим покращенням $CVaR_{10}$ при помірній поступці в $E[\Pi]$. На локації Педдапурам система виявляє стратегії з переходом $CVaR_{10}$ через нуль у позитивну зону. На локаціях Паралахемунді та Венсу інформаційна перевага реалізується в іншій формі: система не виявляє стратегій із суттєво кращим ризик-профілем, але кількісно характеризує ризик кожної доступної опції, що є самостійною діагностичною інформацією, недоступною з точкових рекомендацій. У жодному з чотирьох обчислювальних експериментів запропоновану систему не розглянуто як інструмент, що одночасно переважає порівнювані аналоги за всіма критеріями. Натомість вона слугує засобом

структурування простору прийняття рішень, який в альтернативних підходах залишається неструктурованим

4.2.3 Адаптивність системи до цінових сценаріїв. Порівняння трьох цінових сценаріїв у межах чотирьох локацій дозволяє оцінити функціональну перевагу розробленої системи порівняно з альтернативними підходами: здатність перебудовувати оптимальну дозу у відповідь на зміну ринкової кон'юнктури. Фіксована рекомендація при зміні цін залишається незмінною – вона є горизонтальною лінією у просторі "ціновий сценарій – оптимальна доза". Запропонована система формує функцію, а не константу.

Для ілюстрації цієї властивості в табл. 4.8 подано зсув *Opt.MV* і *Opt.CVaR* доз при переході між ціновими сценаріями для кожної локації. Для компактності наведено лише ті добрива, для яких зсув є найбільш вираженим.

Таблиця 4.8

Зсув оптимальних доз системи між ціновими сценаріями

Локація	Добриво	<i>Opt.MV</i> , низькі (кг/га)	<i>Opt.MV</i> , помірний (кг/га)	<i>Opt.MV</i> , кризовий (кг/га)	<i>Opt.CVaR</i> , низькі (кг/га)	<i>Opt.CVaR</i> , помірний (кг/га)	<i>Opt.CVaR</i> , кризовий (кг/га)
Паралахемунді	Карбамід	500	350	350	250	250	250
Паралахемунді	Аміачна селітра	450	450	450	350	350	350
Педдапурам	Карбамід	50	50	50	50	50	50
Педдапурам	НРК 20-20-20	250	250	250	100	100	100
Венсу	Карбамід	150	150	150	50	50	50
Венсу	НРК 15-15-15	400	400	400	50	50	50

Локація	Добриво	<i>Opt.MV</i> , низькі (кг/га)	<i>Opt.MV</i> , помірний (кг/га)	<i>Opt.MV</i> , кризовий (кг/га)	<i>Opt.CVa</i> <i>R</i> , низькі (кг/га)	<i>Opt.CVa</i> <i>R</i> , помірний (кг/га)	<i>Opt.CVa</i> <i>R</i> , кризовий (кг/га)
Хенань	Карбамід	100	100	100	200	200	200
Хенань	Аміачна селітра	150	150	150	250	250	250

З даних табл. 4.8 видно, що для більшості комбінацій локація–добриво оптимальні дози залишаються стабільними при переході між ціновими сценаріями. Виняток – карбамід на локації Паралахемунді, де *Opt.MV* зсувається з 500 кг/га у сценарії низьких цін до 350 кг/га у помірному і кризовому сценаріях. Це пояснюється тим, що при низьких цінах на азот (5,16 грн/кг) висока доза карбаміду залишається рентабельною навіть при нижчій ціні кукурудзи; при подорожчанні азоту до 7,74 і 10,75 грн/кг за умов того самого агрокліматичного ризику *MV*-критерій зсуває оптимум до нижчої дози.

Водночас стабільність оптимальних доз у більшості рядків таблиці не означає відсутності реакції системи на цінові сценарії. Навіть при незмінній оптимальній дозі *E[Π]* і *CVaR₁₀* у цій точці суттєво змінюються між сценаріями – і саме ці числа, доступні з системи для кожного сценарію, є тією інформацією, якої не надає фіксована рекомендація. Зокрема, на локації Хенань рекомендація Rep при переході від сценарію низьких цін до кризового змінює *CVaR₁₀* з -126,68 до 4 403,08 грн/га, тобто кон'юнктурний ризик у найгірших реалізаціях повністю змінює знак. Фермер, що має доступ до системи, бачить цей зсув явно і може прийняти рішення про коригування дози або страхову стратегію; фермер, що дотримується точкової рекомендації, цього зсуву не спостерігає у структурованому вигляді.

Таким чином, адаптивність системи до цінових сценаріїв полягає не стільки в частих змінах оптимальної дози, скільки у здатності кількісно оцінювати наслідки кожного цінового сценарію для всіх доступних стратегій. Така

аналітична можливість не передбачена в порівнюваних базових моделях, оскільки їхні рекомендації сформовано без урахування варіативності цінової кон'юнктури.

4.2.4 Узагальнення: інформаційна цінність системи. Результати чотирьох обчислювальних циклів дозволяють сформулювати узагальнений висновок щодо характеру переваги запропонованої системи відносно розглянутих аналогів.

У жодному з чотирьох порівнюваних підходів не наведено кількісної оцінки ризику запропонованих рекомендацій. У працях Sairam et al. [50], Sujatha et al. [100], Gao et al. [101] та Ren et al. [102] визначено точкові дози NPK, оптимізовані за агрономічними чи агрономічно-економічними критеріями в детермінованій постановці. За нашими даними, $CVaR_{10}$ цих рекомендацій у відтворених агрокліматичних умовах відповідних локацій становить від -12 296,72 грн/га (Sairam, помірний сценарій) до +2 978,91 грн/га (Ren, помірний сценарій). Ця числова характеристика є прихованою для користувача, який дотримується точкової рекомендації, і стає явною лише через запропоновану систему.

Розроблена система не обмежується розрахунком єдиної рекомендованої дози, а генерує безперервний розподіл техніко-економічних показників з урахуванням ризиків для всього діапазону норм внесення добрив і кожного цінового сценарію. На отриманій кривій відображаються дві орієнтирні точки – $Opt.MV$ і $Opt.CVaR$, що відповідають різним типам ризик-профілю суб'єктів господарювання, залишаючи остаточний вибір безпосередньо за виробником.

Практичним підтвердженням такої інформаційної переваги є результати моделювання для регіону Хенань. У помірному ціновому сценарії фіксована рекомендація Ren забезпечує значення $CVaR_{10} = 2\,978,91$ грн/га тоді як запропонована система на тій самій локації дозволяє визначити ефективнішу стратегію (застосування NPK 15-15-15 або NPK 20-20-20 за еквівалентної дози діючої речовини 30 кг/га для кожного елемента), де $CVaR_{10}$ зростає до 10 318,31 грн/га. Таким чином, мінімізація ризику в найгіршому децилі сценаріїв становить 7 339 грн/га, що досягається за компромісного зниження очікуваного прибутку $E[\Pi]$ на 8 285 грн/га (21,2%). Наявність зазначеної оптимальної точки в просторі

"прибуток–ризик" неможливо виявити без суцільного перебору доз у стохастичній постановці завдання

Аналогічний за структурою результат зафіксовано на локації Педдапурам, де система виявляє стратегії з переходом $CVaR_{10}$ через нуль у позитивну зону при помірній поступці відносно $E[\Pi]$ рекомендації Sujatha. На локаціях Паралахемунді та Венсу такого переходу у дослідженому діапазоні доз не виявлено, однак і там система надає кількісну діагностику рівня ризику. Зокрема, для локації Паралахемунді встановлено, що CV прибутку для рекомендації Sairam становить 78,4%, що є найвищим значенням серед чотирьох локацій і свідчить про найбільшу чутливість економічного результату до погодної мінливості саме в умовах Одіші.

Таким чином, у проведених обчислювальних циклах запропонована система демонструє якісно іншу інформаційну цінність рішення порівняно з аналогами: вона не замінює рекомендацію точкову на іншу точкову, а замінює точку на структурований простір вибору з кількісно охарактеризованими наслідками кожної опції. Адаптивність до цінових сценаріїв, зафіксована у пп. 4.2.3, є додатковою властивістю цього простору, відсутньою у всіх чотирьох розглянутих аналогів.

4.3. Узагальнення результатів обчислювальної апробації та практичні рекомендації щодо застосування системи

Результати проведеної обчислювальної апробації показують, що розроблена гібридна система оптимізації доз мінеральних добрив є працездатною з погляду прогнозної точності та придатною як інструмент підтримки агроекономічних рішень у стохастичних умовах. Поєднання прогнозної моделі врожайності з модулем оптимізації за прибутковістю з урахуванням ризику забезпечує перехід від традиційного підходу, орієнтованого на агрономічні норми внесення, до системи ухвалення рішень, у якій враховуються врожайність,

вартість поживних елементів, ринкова ціна продукції та варіативність зовнішніх умов.

4.3.1 Аналіз результатів тестування запропонованої моделі. Результати тестування прогнозного ядра, описаного у розділі 2, показують прийнятний рівень узагальнювальної здатності на тестовій вибірці. Для тренувальної підвибірки значення MSE становило 74488.04, RMSE – 272.92, MAE – 181.38, MedAE – 114.33, MAPE – 6.62%, sMAPE – 6.55%, WAPE – 6.01%. Для тестової підвибірки відповідні значення дорівнювали 338124.43, 581.48, 376.77, 259.87, 14.41%, 14.10% і 12.42%. Наведені у табл. 4.9 результати показують, що модель зберігає прийнятний рівень точності на незалежних даних, а відносні похибки прогнозу залишаються в межах, придатних для використання моделі в задачах оптимізації за прибутковістю з урахуванням ризику.

Таблиця 4.9

Метрики точності прогнозової моделі

Метрика	Train	Test
MSE	74488.04	338124.4
RMSE	272.92	581.48
MAE	181.38	376.77
MedAE	114.33	259.87
MAPE (%)	6.62	14.41
sMAPE (%)	6.55	14.1
WAPE (%)	6.01	12.42

Похибки оцінювалися як в абсолютних, так і у відносних метриках, що дає змогу інтерпретувати точність моделі у прикладному агроекономічному контексті, де похибка прогнозу безпосередньо трансформується у зміщення оцінки доходу та прибутку.

Порівняльний аналіз із базовими моделями, виконаний у межах апробації, відтворено на тестовій вибірці (табл. 4.10). Так, модель лінійної регресії продемонструвала на тестовій вибірці RMSE 1189.32, MAE 727.68, MAPE

24.42%, sMAPE 23.23%, WAPE 23.99% і коефіцієнт детермінації $R^2=0.6326$. Натомість градієнтний бустинг показав суттєво кращі результати: RMSE 581.48, MAE 376.77, MAPE 14.41%, sMAPE 14.10%, WAPE 12.42%, $R^2=0.9122$. Модель випадкового лісу, у свою чергу, також продемонструвала дуже високі показники: RMSE 575.56, MAE 344.95, MAPE 13.10%, sMAPE 12.87%, WAPE 11.37%, $R^2=0.9140$. Це показує, що у задачі прогнозування врожайності лінійна форма залежності є недостатньою для коректного опису взаємодій між погодними показниками, структурним контекстом країни та вектором доз NPK; обґрунтування вибору ансамблевих моделей подано у розділі 2.

Таблиця 4.10

Порівняння метрик якості для досліджених моделей на навчальній і тестовій вибірках

Модель	Метрика	Train	Test
Лінійна регресія	MSE	1462440	1414494
	RMSE	1209.31	1189.32
	MAE	753.78	727.68
	MedAE	461.45	429.3
	MAPE (%)	26.54	24.42
	sMAPE (%)	24.54	23.23
	WAPE (%)	24.96	23.99
Градієнтний бустинг	MSE	74488.04	338124.4
	RMSE	272.92	581.48
	MAE	181.38	376.77
	MedAE	114.33	259.87
	MAPE (%)	6.62	14.41
	sMAPE (%)	6.55	14.1
	WAPE (%)	6.01	12.42
Випадковий ліс	MSE	50747.73	331271.9
	RMSE	225.27	575.56
	MAE	132.61	344.95
	MedAE	78.09	224.31

Модель	Метрика	Train	Test
	MAPE (%)	4.86	13.1
	sMAPE (%)	4.82	12.87
	WAPE (%)	4.39	11.37

У табл. 4.11 наведено узагальнені значення коефіцієнта детермінації та параметрів σ^2 і $0,5 \cdot \sigma^2$, які відповідають дисперсії залишків логарифмованого прогнозу та її половині — поправці для логнормальної корекції зворотного перетворення прогнозу з логарифмічного простору в початковий масштаб (математичне обґрунтування — у розділі 2). Для лінійної регресії зафіксовано дисперсію залишків 0.0891466767 та поправку 0.0445733384; для градієнтного бустингу – 0.0375397383 та 0.0187698691; для випадкового лісу – 0.0336123622 та 0.0168061811 відповідно. Менші значення цих параметрів для ансамблевих моделей відповідають нижчому рівню залишкової невизначеності, що узгоджується з вищими значеннями R^2 для градієнтного бустингу та випадкового лісу. За тестовими метриками випадковий ліс має незначну числову перевагу над градієнтним бустингом, але ця перевага є несуттєвою для практичних висновків. Ключовим результатом апробації є підтвердження придатності ансамблевих нелінійних регресорів як прогнозного ядра системи оптимізації.

Таблиця 4.11

Порівняння узагальнювальних показників моделей

Модель	RMSE Train	RMSE Test	R2 Train	R2 Test	σ^2	$0.5 \cdot \sigma^2$
Лінійна регресія	1209.31	1189.32	0.6262	0.6326	0.089147	0.044573
Градієнтний бустинг	272.92	581.48	0.981	0.9122	0.03754	0.01877
Випадковий ліс	225.27	575.56	0.987	0.914	0.033612	0.016806

Важливим компонентом апробації стала оцінка поведінки системи за різних економічних сценаріїв, детально розглянутих у пп. 4.2.3 для чотирьох локацій, що досліджувалися в альтернативних працях. Результати показали, що розроблена система функціонує як інструмент адаптивної оптимізації, чутливий до змін

ринкової кон'юнктури. Розроблений підхід використовує прогнозний блок як проміжний компонент інтегрованої системи, у якій ухвалення рішення формується через оптимізацію з урахуванням ризиків. Керованими змінними є окремі компоненти N , P і K , а не їхня агрегована сума; прогноз врожайності формується з використанням логарифмічного перетворення та поправки на логнормальний характер залишків; структурні міждержавні відмінності враховуються через контекстну змінну `CountryTechLevel`; вибір рішення здійснюється на основі сценарної варіативності цін. Математичне обґрунтування цих елементів подано у розділі 2.

Практична значущість результатів полягає в тому, що розроблена система може бути застосована як основа для комп'ютеризованої підтримки агрономічних рішень у господарствах, що працюють в умовах мінливої ринкової та кліматичної кон'юнктури. Її використання є доцільним у ситуаціях, коли необхідно не просто оцінити потенційний урожай, а обрати таку структуру внесення добрив, яка забезпечує прийнятний компроміс між продуктивністю, витратами і стійкістю фінансового результату. Особливо цінним це є для сценаріїв високої цінової волатильності, коли економічна доцільність внесення кожного елемента живлення вже не може визначатися нормативно або інтуїтивно. У такому середовищі система може виступати основою для формування адаптивних рекомендацій, коригування норм удобрення перед сезоном, а також для порівняльного аналізу альтернативних управлінських стратегій.

На підставі проведеної апробації доцільно сформулювати такі практичні рекомендації щодо застосування системи. По-перше, запропоновану модель варто використовувати не як ізольований інструмент прогнозу, а саме як компонент комплексної системи підтримки рішень, де результати, де результати прогнозування безпосередньо переходять у блок економічного оцінювання. По-друге, при практичному впровадженні слід регулярно актуалізувати сценарні параметри цін на добрива і продукцію, оскільки зміна ринкової кон'юнктури може впливати на оптимальні норми не менше, ніж зміна погодних факторів. По-

третє, для забезпечення коректної роботи моделі необхідно підтримувати актуальність вхідних погодних та контекстних даних, оскільки саме якість вхідного простору визначає надійність прогнозного ядра. По-четверте, у прикладних реалізаціях доцільно проводити періодичне повторне навчання ансамблевої моделі на розширеній вибірці нових сезонів, щоб зберігати її адаптивність до змін агротехнологічних та кліматичних умов. Нарешті, при виборі кінцевої моделі для прикладного використання доцільно враховувати не лише абсолютний максимум точності, а й стабільність роботи моделі в алгоритмі оптимізації за прибутковістю з урахуванням ризику, оскільки саме цей критерій є визначальним для реального управлінського застосування.

Отже, проведена апробація прогнозного ядра охоплює оцінку точності на тестовій вибірці, порівняння з базовими алгоритмами та підтвердження придатності для використання в блоці оптимізації за прибутковістю з урахуванням ризику, описаному у пп. 4.2. Отримані метрики характеризують прогнозне ядро як придатне до використання у задачах точного землеробства.

4.3.2 Програмна реалізація інтелектуальної системи підтримки прийняття рішень. Програмний прототип інтелектуальної системи підтримки прийняття рішень щодо внесення мінеральних добрив, детальний опис архітектури якого наведено у розділі 3, реалізує функції прогнозування врожайності та економічного аналізу. У цьому підпункті продемонстровано результати його роботи у трьох режимах: точковий прогноз урожайності, детермінований аналіз вигоди та сценарна ризик-орієнтована оптимізація.

Узагальнену структуру взаємодії компонентів програмного прототипу подано на рис. 4.1.

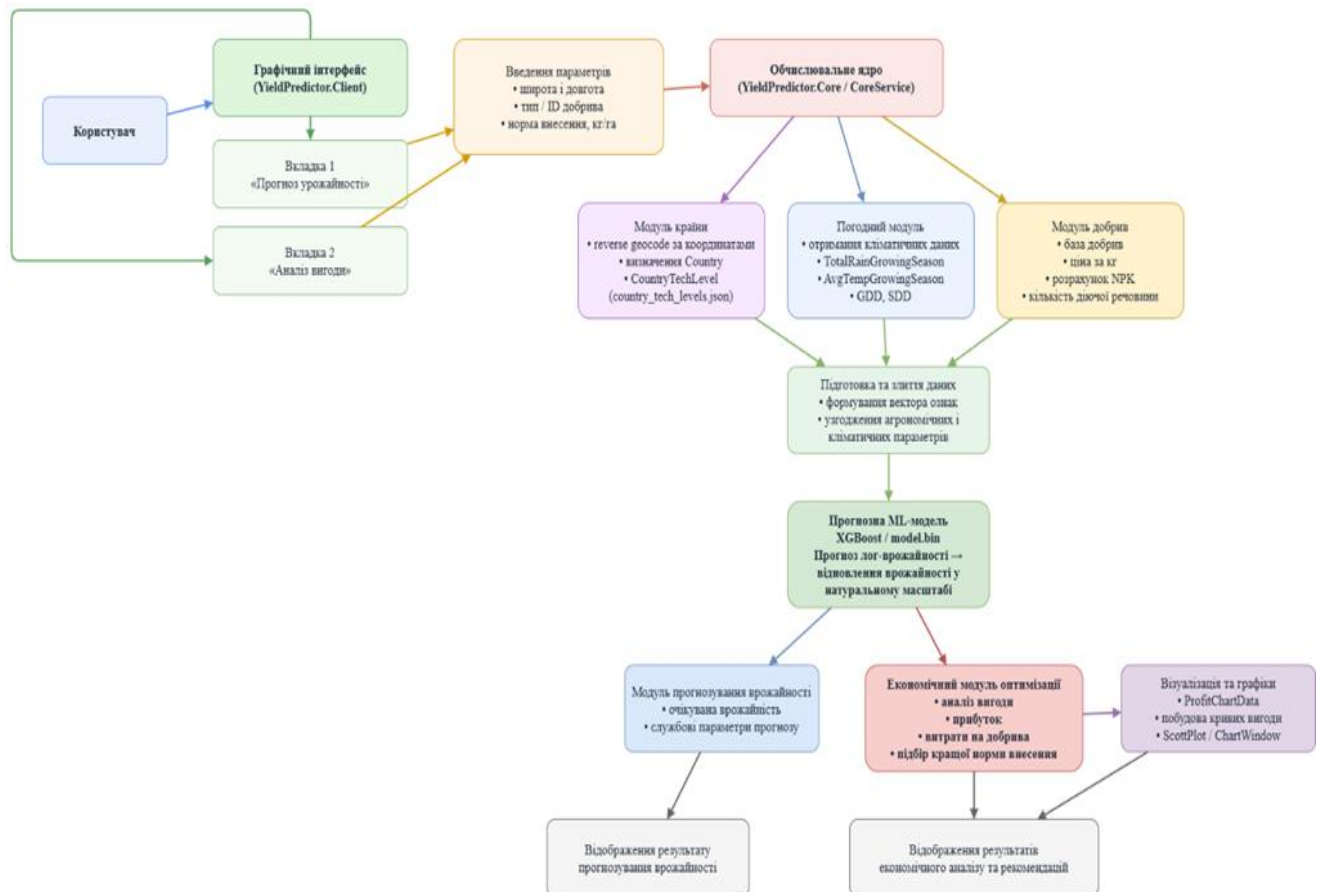


Рис. 4.1 Архітектура програмної системи прогнозування врожайності та оптимізації внесення добрив

Графічний інтерфейс програми реалізовано у вигляді двох основних вкладок: «Прогноз урожайності» та «Аналіз вигоди», що забезпечують логічне розділення функціоналу. Інтерфейс вкладки прогнозування наведено на рис. 4.2.

У цьому модулі користувач вводить такі параметри: тип мінерального добрива; норму внесення добрива та вибирає точку із земельною ділянкою на мапі.

Після введення даних та натискання кнопки «Обчислити» система передає параметри до прогнозної моделі, яка формує оцінку очікуваної врожайності.

Прогноз урожайності

— □ ×

Прогноз урожайності

Аналіз вигоди

Прогноз урожайності

Кількість (кг):

Добриво:

Виберіть точку поля на карті нижче



Координати: натисніть на карту

Обчислити

Точка на мапі не вибрана або неправильна

Рис. 4.2 Інтерфейс модуля прогнозування врожайності

Заповнений приклад інтерфейсу з реальними значеннями параметрів представлено на рис. 4.3.

Прогноз урожайності

Прогноз урожайності

Аналіз вигоди

Сценарний аналіз

Прогноз урожайності

Кількість (кг):

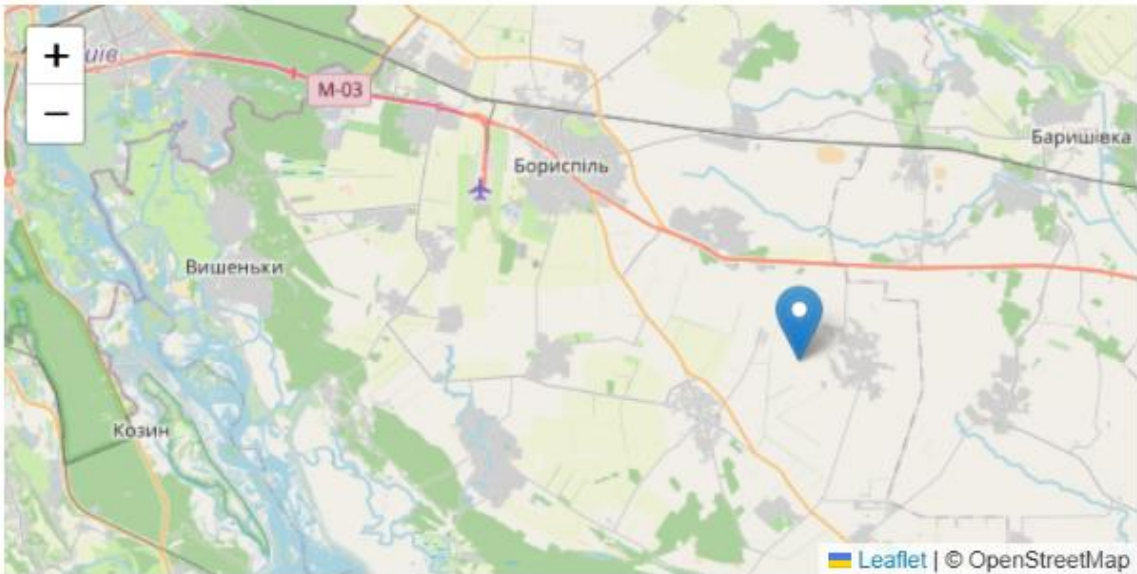
Добриво:

Сульфат калію

Виберіть точку поля на карті нижче

+

—



Координати: 50,2604, 31,1215

Обчислити

Готово до обчислення прогнозу урожайності

Рис. 4.3 Введення параметрів для прогнозування врожайності

Результат роботи моделі відображається у вигляді діалогового вікна (рис. 4.4), де наведено як прогнозне значення врожайності, так і використані вхідні параметри.

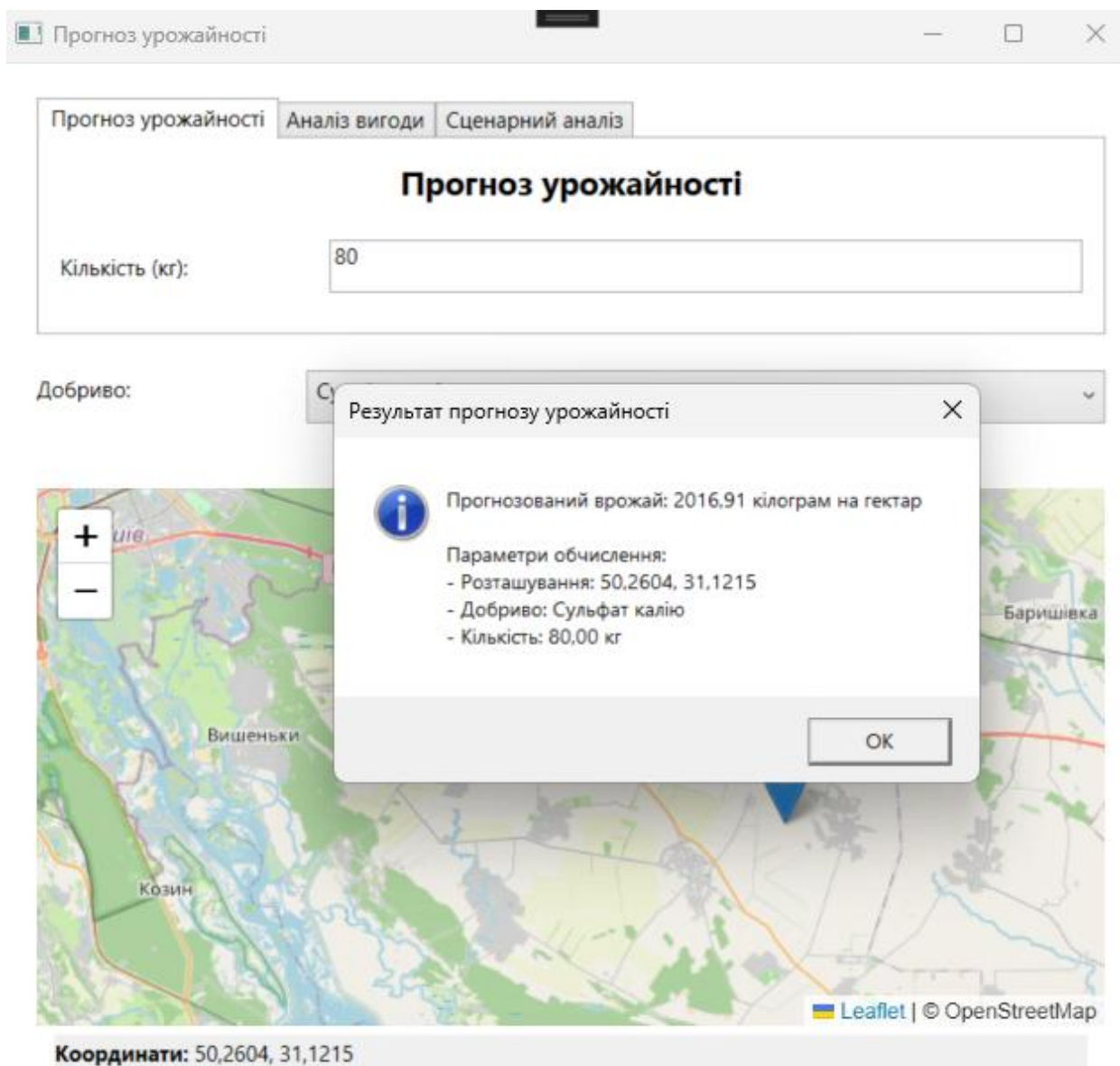


Рис. 4.4 Відображення результату прогнозування врожайності

На рис. 4.4 наведено числовий результат прогнозу разом з використаними вхідними параметрами, що забезпечує прозорість роботи моделі.

Другим функціональним компонентом системи є модуль аналізу вигоди, інтерфейс якого представлено на рис. 4.5.

Прогноз урожайності

Аналіз вигоди

Сценарний аналіз

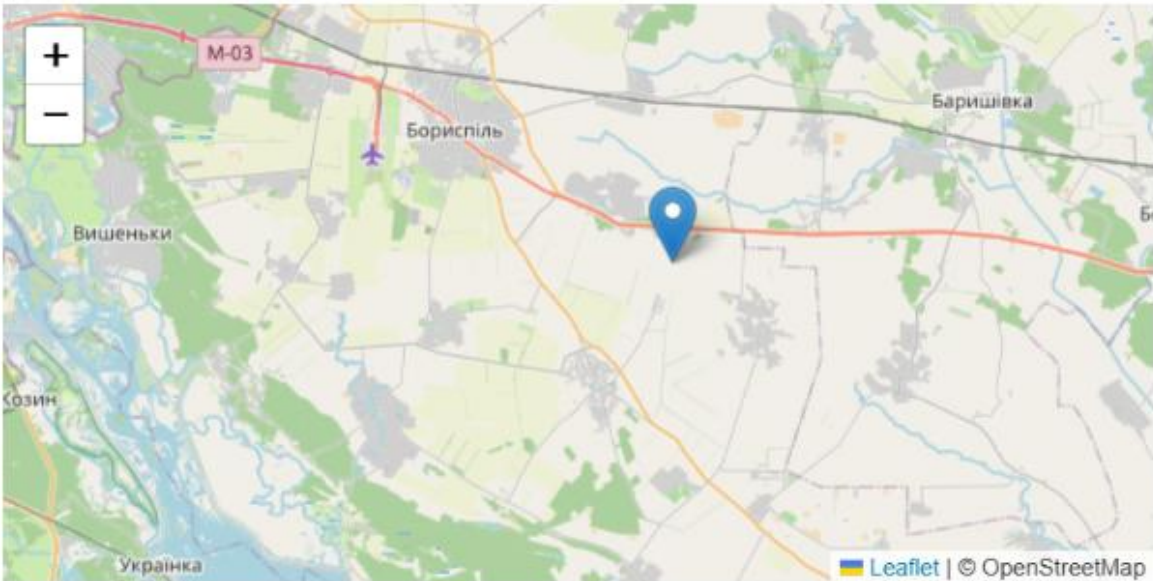
Аналіз вигоди та оптимізація

Ціна урожаю (грн/кг):

Добриво:

Аміачна селітра

Виберіть точку поля на карті нижче



Координати: 50,2902, 31,1050

Побудувати графік

Введіть правильну ціну урожаю (більшу, ніж 0)

Рис. 4.5 Інтерфейс модуля аналізу економічної ефективності

У цьому модулі, окрім вибору точки на мапі та типу добрива, користувач додатково задає ринкову ціну продукції (грн/кг).

На основі цих даних система виконує серію обчислень для різних значень норм внесення добрив та формує залежність прибутку від кількості внесеного ресурсу.

Заповнений інтерфейс перед виконанням обчислень наведено на рис. 4.6.

Прогноз урожайності

Прогноз урожайності Аналіз вигоди Сценарний аналіз

Аналіз вигоди та оптимізація

Ціна урожаю (грн/кг): 9

Добриво: Аміачна селітра

Виберіть точку поля на карті нижче

Координати: 50,2902, 31,1050

Побудувати графік

Готово до побудови графіка аналізу вигоди

Рис. 4.6 Введення параметрів для економічного аналізу

Результатом роботи модуля є графічна інтерпретація залежності прибутку від норми внесення добрив, яка представлена на рис. 4.7.

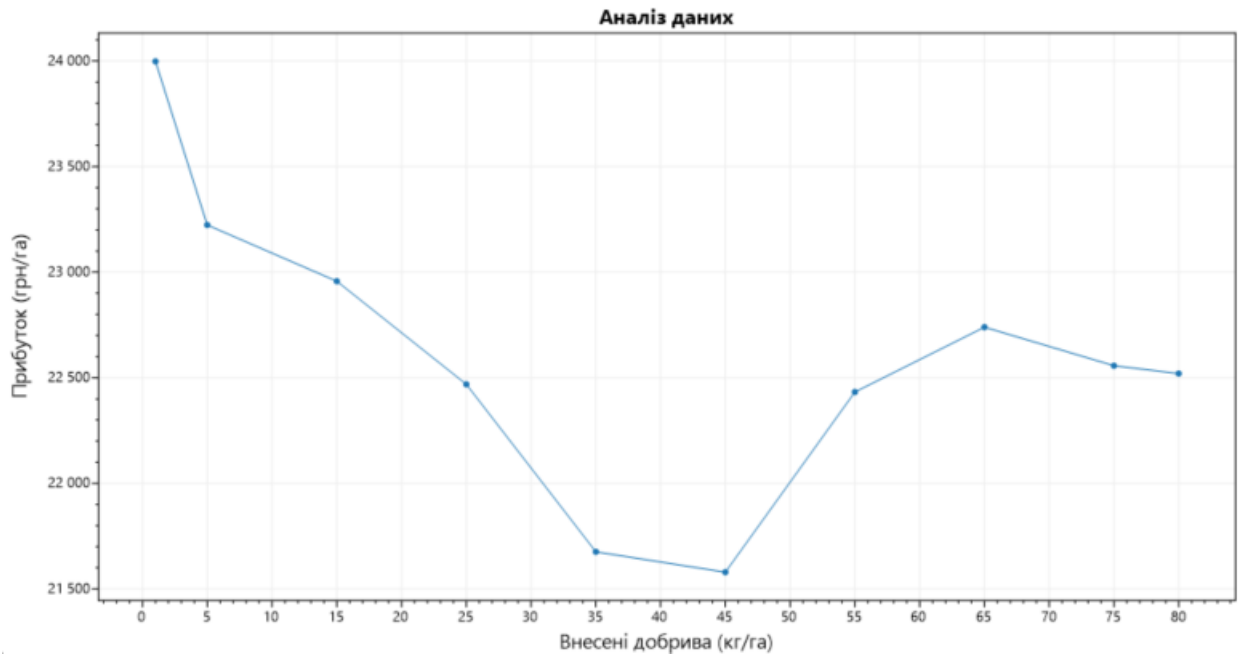


Рис. 4.7 Графік залежності прибутку від норми внесення добрив

Аналіз отриманого графіка свідчить про наявність нелінійної залежності між обсягом внесення добрив і фінансовим результатом. Зокрема, спостерігається зростання прибутку на початкових етапах збільшення норми внесення, після чого досягається максимум, а подальше збільшення кількості добрив призводить до зниження економічної ефективності. Така поведінка підтверджує доцільність використання оптимізаційного підходу, оскільки максимізація врожайності не еквівалентна максимізації прибутку.

З функціональної точки зору алгоритм роботи системи можна описати таким чином. На першому етапі користувач вводить вхідні параметри, які характеризують агрокліматичні умови та структуру внесення добрив. Далі ці параметри проходять попередню обробку та трансформацію відповідно до вимог навченої моделі. Після цього прогнозна модель обчислює очікувану врожайність із урахуванням нелінійних залежностей між факторами.

У випадку використання модуля економічного аналізу отримане значення врожайності використовується для обчислення доходу, після чого з урахуванням вартості добрив визначається прибуток. Далі система виконує варіювання норми

внесення та формує функцію прибутку, що дозволяє визначити оптимальне значення.

Окрім детермінованої оптимізації, описаної вище, програмний прототип реалізує функцію сценарної оптимізації за прибутковістю з урахуванням ризику, призначену для кількісного оцінювання ризикових характеристик управлінських рішень. Принципи її роботи, формальна постановка та алгоритмічні деталі викладено в підрозділах 2.1.2 та 3.3.6. У цьому підпункті продемонстровано результати її застосування на тестовій локації Київщини (на південний схід від Борисполя) при ринковій ціні пшениці 9 грн/кг, фіксованих операційних витратах 10 000 грн/га та параметрі ризик-аверсії $\lambda = 1.0$. Тип внесеного добрива – аміачна селітра з масовою часткою азоту 34%. Інтерфейс модуля та отримані результати наведено на рис. 4.8.

Аналіз отриманих метрик ілюструє ключову методологічну особливість запропонованого підходу – здатність системи генерувати дві альтернативні рекомендації залежно від обраного критерію оптимальності. Стратегія Mean-Variance ідентифікує оптимальною дозу 55 кг/га з очікуваним прибутком 9956.57 грн/га, що відповідає балансу середнього економічного результату та стабільності у заданому ризиковому профілі. Стратегія максимізації CVaR₁₀ ідентифікує оптимальною дозу 45 кг/га з очікуваним прибутком 9281.64 грн/га та CVaR₁₀ = -9567.72 грн/га, що відповідає мінімізації вразливості до найгірших погодних реалізацій. Обидві стратегії узгоджуються з нормативним діапазоном внесення азотних добрив для озимої пшениці в умовах України (типово 30-90 кг/га діючої речовини), однак Mean-Variance критерій надає перевагу дещо більшій дозі задля компенсації варіативності прогнозу врожайності.

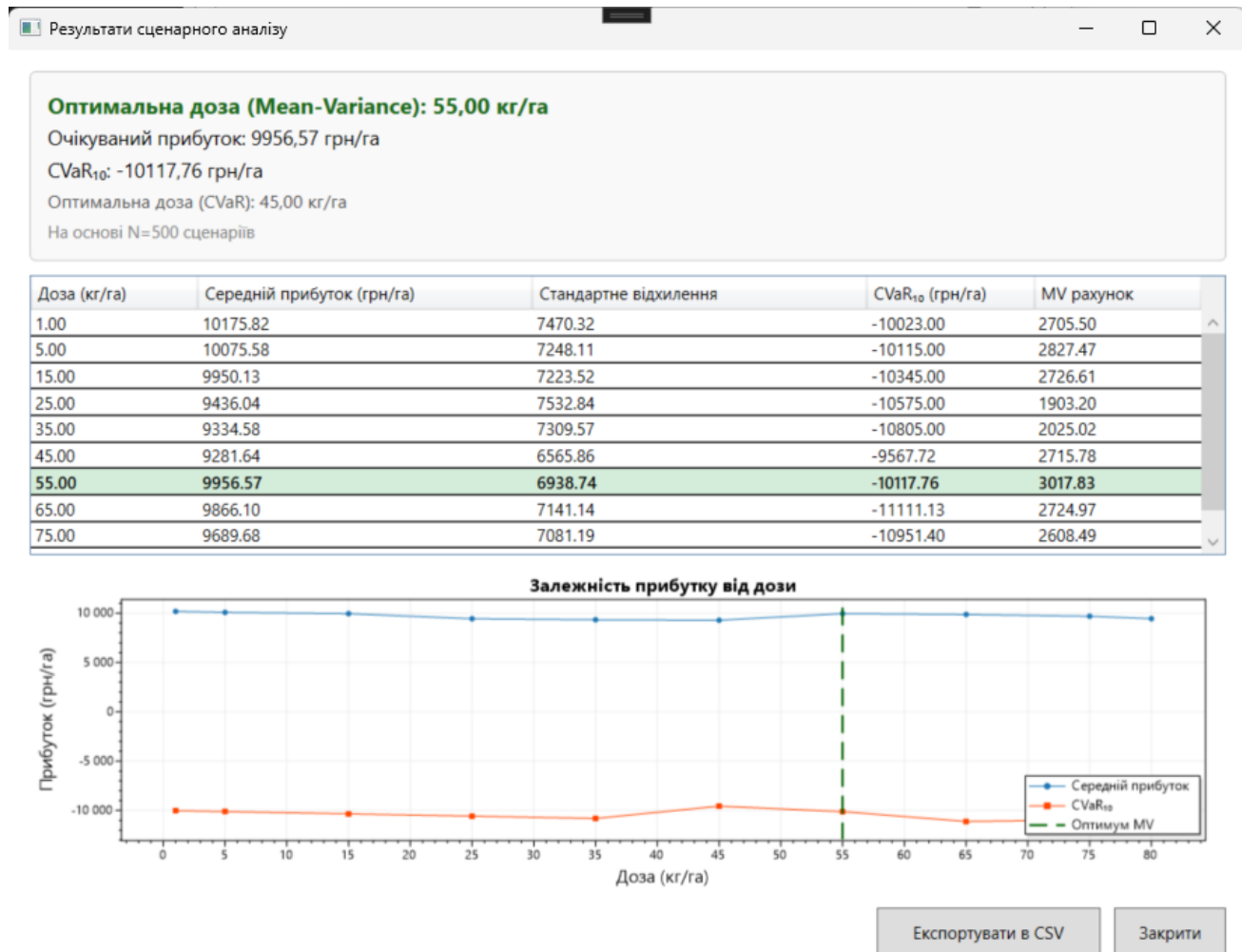


Рис. 4.8 Результати сценарної оптимізації за прибутковістю з урахуванням ризику для тестової локації Київщини

Від'ємні значення $CVaR_{10}$ для всіх доз показують відчутний рівень ризику виробництва пшениці в досліджуваній локації за наявних цінових параметрів. Цей діагностичний результат недоступний з традиційних детермінованих рекомендацій: розроблена система кількісно ідентифікує залишковий ризик і обґрунтовує доцільність застосування додаткових механізмів управління — страхування врожаю, хеджування цінових ризиків, диверсифікації культур або вибору стійкіших сортів. Сценарний модуль виконує дві функції — оптимізацію за обраним критерієм та діагностику фінансової стійкості агровиробництва.

Програмна реалізація демонструє роботу системи на конкретному прикладі та її придатність до використання у задачах підтримки прийняття рішень в аграрному секторі.

Висновки до розділу 4

Підтверджено високу прогностичну точність розробленої гібридної системи. На тестовій вибірці ансамблева модель градієнтного бустингу (XGBoost) досягла коефіцієнта детермінації $R^2 = 0.9122$, $RMSE = 581.48$ кг/га та $MAE = 376.77$ кг/га, що значно перевищує результати базової лінійної регресії $R^2 = 0.6326$, $RMSE = 1189.32$ кг/га). Це доводить здатність системи адекватно апроксимувати нелінійний відгук урожайності на дози добрив і погодні фактори.

Верифікацію прогнозного ядра проведено шляхом порівняння прогнозів моделі з дозами внесення мінеральних добрив, рекомендованими у роботах інших авторів на тих самих географічних локаціях, де ці дози було отримано. У трьох обчислювальних циклах для кукурудзи прогноз моделі систематично нижчий за повідомлену в роботах врожайність на 52–55%. Для однієї з локацій відхилення становить 6,0%. Виявлене заниження є очікуваним наслідком навчання моделі на агрегованих міжкраїнних даних FAOSTAT і характеризує її як інструмент крос-національного прогнозу в умовах середньої агротехнологічної інтенсивності, а не як заміну польовому експерименту з оптимально керованими умовами.

Розширений порівняльний аналіз у пп. 4.2 показав, що цінність розробленої системи проявляється не у вищому очікуваному прибутку, а в наявності інформації, недоступної з точкових рекомендацій інших авторів. В альтернативних підходах подається лише точкова рекомендована доза без кількісної оцінки ризику; розроблена система генерує повну криву ризик-економічних показників для всього діапазону доз з виділеними точками *Opt.MV* і *Opt.CVaR*. На локації Хенань зафіксовано стратегію NPK 15-15-15 при дозі 200 кг/га, для якої $CVaR_{10}$ у помірному ціновому сценарії становить 10 318 грн/га, що у 3,46 рази перевищує $CVaR_{10}$ фіксованої рекомендації Ren (2 979 грн/га) при

поступці в очікуваному прибутку 21,2%. На локації Педдапурам виявлено стратегії з переходом $CVaR_{10}$ з від'ємних у додатні значення. У всіх чотирьох обчислювальних циклах кількісно охарактеризовано рівень ризику, що відсутній у точкових рекомендаціях. Додатковим результатом є чутливість системи до змін цінової кон'юнктури в межах трьох цінових сценаріїв.

Практичну придатність методології підтверджено розробкою програмного прототипу. Апробація системи на тестовій локації довела здатність алгоритму формувати диференційовані рекомендації залежно від обраного критерію: максимізації балансу «прибуток-ризик» або мінімізації екстремальних втрат. Це перетворює систему на дієвий механізм управління фінансовою стійкістю агровиробництва.

ВИСНОВКИ

У дисертаційній роботі розв’язано актуальну науково-прикладну задачу розроблення інтелектуальних інформаційних технологій в системі управління внесенням мінеральних добрив у стохастичних погодних умовах, що забезпечує автоматизовану підтримку прийняття агрономічних рішень з кількісним урахуванням кліматичних та цінових ризиків.

У ході роботи виконано такі теоретичні дослідження та обчислювальну апробацію.

1. Аналіз предметної області засвідчив, що управління мінеральним живленням рослин є однією з найбільш капіталомістких операцій сучасного рослинництва, економічний результат якої одночасно нелінійний за агрохімічними параметрами та стохастичний за погодними умовами. Системний аналіз сучасних підходів до прийняття рішень в агрономії дозволив виявити системну суперечність: значний прогрес у засобах збору даних супроводжується відставанням алгоритмічного апарату прийняття рішень, що залишається переважно детермінованим та ігнорує стохастичну природу погодних ризиків. За результатом аналізу обґрунтовано методологічний розрив між наявністю точних прогнозних моделей та здатністю трансформувати ці прогнози в економічно стійкі управлінські рішення, що визначило напрям подальших досліджень.

2. Сформовано інтегрований інформаційний простір на основі відкритих джерел FAOSTAT та Open-Meteo Archive API за період 1990–2023 років для близько 180 країн, що після квантильної фільтрації містить 5717 спостережень у форматі «країна–рік–агрегований показник». Розроблено власну структуру вхідного простору ознак, що включає вектор діючих речовин $u = (N, P, K)$, сезонні агрокліматичні індикатори: накопичену суму активних температур GDD та індекс теплового стресу SDD з широтно-залежним порогом, а також контекстний регресор CountryTechLevel як проксі сукупної продуктивності агровиробничої системи країни.

3. Розроблено прогнозну модель врожайності на основі ансамблевого методу градієнтного бустингу XGBoost у логарифмічному просторі цільової змінної з аналітичною корекцією логнормального зміщення через множник $\exp(\sigma^2/2)$, що забезпечує незміщеність математичного сподівання економічних оцінок. Верифікація моделі на тестовій вибірці продемонструвала $R^2 = 0.9122$, $RMSE = 581.48$ кг/га та $MAE = 376.77$ кг/га; п'ятикратна крос-валідація ($R^2 = 0.908 \pm 0.010$, $RMSE = 599 \pm 24$ кг/га) підтвердила стабільність узагальнювальної здатності. Формалізований абляційний аналіз структури вхідного простору ознак показав, що ознака CountryTechLevel забезпечує приріст коефіцієнта детермінації з 0.807 до 0.9122 (+0.104), тоді як альтернативний трендовий регресор (календарний рік) дає значно менший приріст (+0.020) і свідомо відхилений як методологічно некоректний для задач ex ante прогнозування. Порівняльний аналіз з класичною лінійною регресією ($RMSE = 1189.32$ кг/га, $MAE = 727.68$ кг/га, $R^2 = 0.6326$) засвідчив суттєву перевагу нелінійних ансамблевих підходів, що пояснюється нездатністю лінійної форми відтворити фундаментальний агрономічний закон спадної віддачі.

4. Розроблено математичну модель оптимізації внесення добрив за прибутковістю з урахуванням ризику, що інтегрує нелінійну прогнозну функцію відгуку врожайності з ризик-орієнтованим економічним функціоналом на основі поєднання Mean-Variance скаляризації та емпірично обчисленої метрики Conditional Value-at-Risk (CVaR) на сценарному просторі погодних і цінових реалізацій. Запропонована постановка забезпечує перехід від точкових агрономічних рекомендацій до повноцінного управління ризиками, формуючи стратегії, стійкі до несприятливих погодних і цінових сценаріїв.

5. Запропоновано та програмно реалізовано інформаційну технологію у вигляді двокomпонентної архітектури програмного комплексу, що архітектурно відокремлює дослідницьке середовище підготовки даних і машинного навчання (Python) від модуля бізнес-логіки клієнтського застосунку, який відповідає за агроекономічні розрахунки, сценарний аналіз та інтерфейс користувача (.NET). Це

досягається через серіалізацію прогнозної моделі у платформонезалежний формат. Така архітектурна організація знімає типову для наукових прототипів проблему розриву між середовищем розробки та прикладним застосуванням, забезпечуючи інтерактивну швидкодію модуля сценарної оптимізації на рівні 1.57 ± 0.05 мс на повний цикл перебору без міжпроцесної взаємодії з ML-бекендом. Реалізовано модульну структуру з виділенням бібліотеки ядра, консольної утиліти пакетного збору даних та клієнтського застосунку з картографічним інтерфейсом. Розроблено механізми відмовостійкої взаємодії з відкритими API: повторні спроби з фіксованою паузою 60 с при HTTP 429 та лінійним збільшенням затримки при інших помилках, проміжне резервне копіювання при пакетному зборі даних, що забезпечує надійність роботи в умовах нестабільного мережевого середовища.

6. Обчислювальна апробація розробленої технології, проведена для чотирьох географічних локацій в Індії та Китаї, підтвердила, що запропонована система забезпечує вищу інформаційну цінність рішень порівняно з існуючими аналогами. На основі моделювання стохастичної множини сценаріїв погодних умов та трьох цінових ситуацій визначено, що традиційні підходи обмежуються формуванням точкової рекомендованої дози NPK без кількісного оцінювання ризиків. Натомість розроблена система генерує безперервну криву залежності економічних показників від ризику в межах усього діапазону доз із виділенням характерних точок *Opt.MV* та *Opt.CVaR*. Зокрема, для локації Хенань ідентифіковано стратегію застосування комплексного добрива NPK 15-15-15 у дозі 200 кг/га, за якої показник *CVaR₁₀* у помірному ціновому сценарії становить 10 318 грн/га, що у 3,46 рази перевищує рівень ризику для традиційної фіксованої рекомендації (2 979 грн/га) за компромісного зниження очікуваного прибутку на 21,2%. Для локації Педдапурам система дозволяє виявити ефективні стратегії, що забезпечують перехід показника *CVaR₁₀* від від'ємних до додатних значень, що є неможливим за використання статичних рекомендацій.

7. У межах трьох досліджених цінових сценаріїв (низького, помірного та кризового) підтверджено високу адаптивність розробленої системи до змін

ринкової кон'юнктури. Для кожного сценарію система кількісно оцінює математичне сподівання прибутку, показник $CVaR_{10}$ та коефіцієнт варіації для всіх точок діапазону норм внесення добрив. Це створює надійну інформаційну основу для вибору стратегії управління мінеральним живленням відповідно до індивідуального ризик-профілю особи, яка приймає рішення. Зазначена аналітична можливість не реалізована у фіксованих рекомендаціях, які залишаються статичними за будь-яких варіацій цін. Таким чином, створена технологія функціонує як динамічний інструмент структурованого вибору компромісних рішень між рівнем очікуваного прибутку та ступенем фінансового ризику в умовах мінливості зовнішнього економічного середовища.

8. Запропоновано удосконалену архітектуру систему підтримки прийняття рішень для управління внесенням добрив, яка поєднує три ключові компоненти:

- 1) методи машинного навчання;
- 2) модель оптимізації за прибутковістю з урахуванням ризику;
- 3) сучасні програмно-інженерні рішення.

Розроблений програмний прототип демонструє працездатність запропонованої методології наскрізного аналітичного процесу «збір даних з відкритих джерел → формування простору ознак → прогноз врожайності → оптимізація внесення добрив за прибутковістю з урахуванням ризику → ризик-орієнтована рекомендація», та може слугувати основою для подальшого виробничого впровадження у системах точного землеробства. Подальший розвиток прототипу до повноцінного виробничого рішення передбачає інтеграцію ґрунтової складової простору ознак (агрохімічний аналіз, дані локального моніторингу), підключення до зовнішніх джерел актуальних ринкових цін на добрива та продукцію через відповідні API, а також адаптацію моделі до рівня окремого господарства методами transfer learning.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. AI Tool to Help Farmers Measure Real-Time Crop Health from the Field. *Global AgTech Initiative*. URL: <https://www.globalagtechinitiative.com/digital-farming/analytics/ai-tool-to-help-farmers-measure-real-time-crop-health-from-the-field/>.
2. Cai B., Shi F., Geremew B. A., Addis A. K., Abate M. C., Dessie W., Bayu T. Precision agriculture techniques for optimizing chemical fertilizer use and environmental sustainability: a systematic review. *Frontiers in Agronomy*. 2025. Vol. 7. Article 1665444. DOI: 10.3389/fagro.2025.1665444.
3. A Systematic Review of IoT Solutions for Smart Farming. *MDPI Sensors*. 2020. Vol. 20, No. 15. Article 4231. URL: <https://www.mdpi.com/1424-8220/20/15/4231>.
4. AI in agriculture: pros, cons and how to stay ahead. *BPM Insights*. URL: <https://www.bpm.com/insights/ai-in-agriculture/>.
5. Lysenko V., Bolbot I., Romasevych Y., Loveykin V., Voytiuk V. Algorithms of robotic electrotechnical complex control in agricultural production. *Robotic Electrotechnical Complexes in Agricultural Engineering*. 2023. Chapter 13. DOI: 10.1201/9781003337706-13.
6. Dolia M., Lysenko V., Lendiel T., Nakonechna K., Humeniuk L. Neuron network prediction of damage of E. integriceps bug on winter wheat in Ukraine. *Scientific Reports of the National University of Life and Environmental Sciences of Ukraine*. 2024. Vol. 20 (4). P. 96–105. DOI: 10.31548/dopovid/3.2024.96.
7. Lutska N., Vlasenko L., Zaiets N., Lysenko V. Modeling the Productivity of a Sugar Factory using Machine Learning Methods. 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT). 2022. P. 353–356. DOI: 10.1109/CSIT56902.2022.10000571.
8. Cloud–edge–device collaborative computing in smart agriculture: architectures, applications, and future perspectives. *PubMed Central*. 2025. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC12558850/>.

9. Variable Rate Technology and Its Application in Precision Agriculture. *UF/IFAS Extension*. Publication AE607. URL: <https://edis.ifas.ufl.edu/publication/AE607>.
10. Variable-Rate Fertilization for Summer Maize Using Combined Proximal Sensing Technology and the Nitrogen Balance Principle. *Agriculture (MDPI)*. 2024. Vol. 14, No. 7. Article 1180. URL: <https://www.mdpi.com/2077-0472/14/7/1180>.
11. Chen X., Zhang H., Wong C. U. I. Dynamic monitoring and precision fertilization decision system for agricultural soil nutrients using UAV remote sensing and GIS. *Agriculture (MDPI)*. 2025. Vol. 15, No. 15. P. 1627. DOI: 10.3390/agriculture15151627.
12. AgriFusion: An Architecture For IoT And Emerging Technologies Based On A Precision Agriculture Survey. *IEEE Xplore*. URL: <https://ieeexplore.ieee.org/iel7/6287639/6514899/09552863.pdf>.
13. Development and Evaluation of an Affordable Variable Rate Applicator Controller for Precision Agriculture. *AgriEngineering (MDPI)*. 2024. Vol. 6, No. 4. Article 265. URL: <https://www.mdpi.com/2624-7402/6/4/265>.
14. Machine learning versus crop growth models: an ally, not a rival. *AoB PLANTS*. 2023. Vol. 15, No. 2. Article plac061. URL: <https://academic.oup.com/aobpla/article/15/2/plac061/6855705>.
15. A review of innovative fertilization strategies in precision agriculture. *Open Access Research Journal of Science and Technology*. 2024. URL: <https://oarjst.com/sites/default/files/OARJST-2024-0137.pdf>.
16. Evaluation of an autonomous smart system for optimal management of fertigation with variable sources of irrigation water. *PubMed Central*. 2023. URL: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10130640/>.
17. Recent Advances in Concept Drift Adaptation Methods for Deep Learning. *Proceedings of IJCAI 2022*. 2022. URL: <https://www.ijcai.org/proceedings/2022/788>.
18. Matchmaker: Data Drift Mitigation in Machine Learning for Large-Scale Systems. *Proceedings of MLSys 2022*. Microsoft Research. URL:

<https://www.microsoft.com/en-us/research/wp-content/uploads/2022/01/MLSYS2022.pdf>.

19. Concept Drift vs Data Drift: Why It Matters in AI. *Viso.ai*. URL: <https://viso.ai/deep-learning/concept-drift-vs-data-drift/>.

20. Transfer Learning In Agriculture. *Meegle Topics*. URL: https://www.meegle.com/en_us/topics/transfer-learning/transfer-learning-in-agriculture.

21. Cetegen E. Optimal design of controlled environment agricultural systems under market uncertainty: PhD dissertation. University of Connecticut, 2021. URL: <https://psor.uconn.edu/wp-content/uploads/sites/1972/2021/11/optimal-design-of-controlled-environment-agriculture-Cetegen-2021.pdf>.

22. Decision Making Under Uncertainties for Adapting to Climate Change in Project Environmental Assessments. Environment Canada, 2011. URL: https://publications.gc.ca/collections/collection_2011/ec/En106-98-2011-eng.pdf.

23. Wu J., et al. Optimizing nitrogen management with deep reinforcement learning and crop simulations. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW 2022), AgriVision Workshop*. 2022. URL: https://openaccess.thecvf.com/content/CVPR2022W/AgriVision/papers/Wu_Optimizing_Nitrogen_Management_With_Deep_Reinforcement_Learning_and_Crop_Simulation_s_CVPRW_2022_paper.pdf.

24. Musanase C., Vodacek A., Hanyurwimfura D., Uwitonze A., Kabandana I. Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices. *Agriculture (MDPI)*. 2023. Vol. 13, No. 11. Article 2141. URL: <https://www.mdpi.com/2077-0472/13/11/2141>.

25. Hasan M., et al. Ensemble machine learning-based recommendation. *Frontiers in Plant Science*. 2023. URL: <https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2023.1234555/full>.

26. Mia M. S., Habibi L. N., Matsui T., Tanaka T. S. T. A Bayesian approach to assessing uncertainty in the effect of fertilization strategies on paddy rice yield via

multiple on-farm experiments. *Plant Production Science*. 2024. DOI: 10.1080/1343943X.2024.2367578.

27. Venkateswara S. M., et al. Interpretable deep learning models for independent fertilizer/crop recommendation. *Scientific Reports*. 2025. Vol. 15. Article 26910. URL: <https://www.nature.com/articles/s41598-025-26910-4>.

28. Wang Z., et al. Reinforcement learning-based agricultural fertilization decision framework. *Drones (MDPI)*. 2025. Vol. 7, No. 8. Article 252. URL: <https://www.mdpi.com/2624-7402/7/8/252>.

29. Parent L. E., et al. Compositional and machine learning tools to model plant nutrient management and fertilizer recommendations. *Horticulturae (MDPI)*. 2025. Vol. 11, No. 2. Article 161. URL: <https://www.mdpi.com/2311-7524/11/2/161>.

30. León Chilito E. D., et al. Sustainability-driven fertilizer recommender system for coffee. *Frontiers in Sustainable Food Systems*. 2025. URL: <https://www.frontiersin.org/journals/sustainable-food-systems/articles/10.3389/fsufs.2024.1445795/pdf>.

31. Baum M. E., Sawyer J. E., Nafziger E. D., et al. The optimum nitrogen fertilizer rate for maize in the US Midwest is increasing. *Nature Communications*. 2025. Vol. 16. Article 404. DOI: 10.1038/s41467-024-55314-7.

32. Manore D., Kelbo H., Abdella O., Abuye F. Optimizing NPSB fertilizer rates for enhanced yield and yield components of mung bean (*Vigna radiata* L.) varieties. *PLOS ONE*. 2025. Vol. 20, No. 10. Article e0321243. DOI: 10.1371/journal.pone.0321243.

33. Lee S. W., Swinton S. M., Basso B. Comparing profitability of variable rate nitrogen prescriptions. *Precision Agriculture*. 2025. Vol. 26. Article 58. DOI: 10.1007/s11119-025-10256-y.

34. Bernard B., Riegert D., Abdella K., Narine S. A Simulated Weather-Driven Bio-Economic Optimization Model for Agricultural Planning. *Mathematics (MDPI)*. 2025. Vol. 13. Article 4010. DOI: 10.3390/math13244010.

35. Kumar K., Parihar C. M., Nayak H. S., et al. Modeling maize growth and nitrogen dynamics using CERES-Maize (DSSAT) under diverse nitrogen management options in a conservation agriculture-based maize-wheat system. *Scientific Reports*. 2024. Vol. 14. Article 11743. DOI: 10.1038/s41598-024-61976-6.
36. Batool M., Sarrazin F. J., Zhang X., et al. Scenario analysis of nitrogen surplus typologies in Europe shows that a 20% fertilizer reduction may fall short of 2030 EU Green Deal goals. *Nature Food*. 2025. Vol. 6. P. 787–798. DOI: 10.1038/s43016-025-01210-2.
37. Nayak H. S., Mkondiwa M., Patra K., Sarkar A., Reddy K. S., Kumar P., Bharadwaj S., Singh Rajbir, Parihar C. M. Risk-return trade-offs in diversified cropping systems under conservation agriculture: Evidence from a 14-year long-term field experiment in north-western India. *European Journal of Agronomy*. 2025. Vol. 162. Article 127436. DOI: 10.1016/j.eja.2024.127436.
38. Li Y., Cao W. Green and Efficient Technology Investment Strategies for a Contract Farming Supply Chain Under the CVaR Criterion. *Sustainability (MDPI)*. 2025. Vol. 17. Article 7600. DOI: 10.3390/su17177600.
39. Mandrini G., Davidson E. A., Nafziger E. D., et al. A market-based insurance approach aligns environmental and economic outcomes in maize nitrogen management. *Communications Earth & Environment*. 2026. Vol. 7. Article 26. DOI: 10.1038/s43247-025-03008-3.
40. Jalilian A., Ghasemi M. M., Ghaznavi S., Moghadam M. S. K., Mobtaker H. G., Dastbala G., Kaab A. Evaluating the effectiveness of variable rate technology of herbicides and chemical fertilizers on reducing environmental pollutants, economic costs and energy use in wheat production. *Results in Engineering*. 2026. Vol. 29. Article 108722. DOI: 10.1016/j.rineng.2025.108722.
41. Paccioretti P., Puntel L., Córdoba M., Mieno T., Ferguson R., Luck J., Thompson L., Balboa G. Site-specific drivers of sensor-based nitrogen management in on-farm corn and wheat experiments. *Frontiers in Agronomy*. 2025. Vol. 7. Article 1651522. DOI: 10.3389/fagro.2025.1651522.

42. Wang L., Gao J., Qureshi W. A. Evolution and Application of Precision Fertilizer: A Review. *Agronomy (MDPI)*. 2025. Vol. 15. Article 1939. DOI: 10.3390/agronomy15081939.
43. Fertilizers by Nutrient. *FAOSTAT, Food and Agriculture Organization of the United Nations*. URL: <https://www.fao.org/faostat/en/#data/RFN>.
44. Fertilizers by Product. *FAOSTAT, Food and Agriculture Organization of the United Nations*. URL: <https://www.fao.org/faostat/en/#data/RFB>.
45. NASA POWER Data Access Viewer (DAV). *Prediction Of Worldwide Energy Resources (POWER)*. URL: <https://power.larc.nasa.gov/data-access-viewer/>.
46. SoilGrids web portal. URL: <https://soilgrids.org/>.
47. Crops and livestock products. *FAOSTAT, Food and Agriculture Organization of the United Nations*. URL: <https://www.fao.org/faostat/en/#data/QCL>.
48. Producer Prices. *FAOSTAT, Food and Agriculture Organization of the United Nations*. URL: <https://www.fao.org/faostat/en/#data/PP>.
49. Free Open-Source Weather API. *Open-Meteo*. URL: <https://open-meteo.com/>.
50. Sairam M., Maitra S., Sagar L., Biswas T., Bárek V., Brestic M., Hossain A. Application of precision nutrient tools for the optimization of fertilizer requirements and assessment of the growth and productivity of maize (*Zea mays* L.) in the northeastern Ghat of India. *Journal of Agriculture and Food Research*. 2025. Vol. 21. Article 101958. DOI: 10.1016/j.jafr.2025.101958.
51. de Lara, A., Mieno, T., Luck, J. D., & Puntel, L. A. (2023). Predicting site-specific economic optimal nitrogen rate using machine learning methods and on-farm precision experimentation. *Precision Agriculture*, 24(5), 1792–1812. <https://doi.org/10.1007/s11119-023-10018-8>
52. Silva F. M., Queirós C., Pereira M., Pinho T., Barroso T., Magalhães S., Boaventura J., Santos F., Cunha M., Martins R. C. Precision Fertilization: A critical review analysis on sensing technologies for nitrogen, phosphorous and potassium

quantification. *Computers and Electronics in Agriculture*. 2024. Vol. 224. Article 109220. DOI: 10.1016/j.compag.2024.109220.

53. Ennaji O., Vergütz L., El Allali A. Machine learning in nutrient management: A review. *Artificial Intelligence in Agriculture*. 2023. Vol. 9. P. 1–11. DOI: 10.1016/j.aiia.2023.06.001.

54. Waqas M., Naseem A., Humphries U. W., Hlaing P. T., Dechpichai P., Wangwongchai A. Applications of machine learning and deep learning in agriculture: A comprehensive review. *Green Technologies and Sustainability*. 2025. Vol. 3, No. 3. Article 100199. DOI: 10.1016/j.grets.2025.100199.

55. Robert P. C. Precision agriculture: a challenge for crop nutrition management. *Plant and Soil*. 2002. Vol. 247. P. 143-149.

56. Pierce F. J., Nowak P. Aspects of Precision Agriculture. *Advances in Agronomy*. 1999. Vol. 67. P. 1-85.

57. Aitchison J., Brown J. A. C. *The Lognormal Distribution*. Cambridge University Press, 1957.

58. Limpert E., Stahel W. A., Abbt M. Log-normal Distributions across the Sciences: Keys and Clues. *BioScience*. 2001. Vol. 51, No. 5. P. 341-352.

59. Rockafellar R. T., Uryasev S. Optimization of Conditional Value-at-Risk. *Journal of Risk*. 2000. Vol. 2, No. 3. P. 21-41.

60. Rockafellar R. T., Uryasev S. Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking & Finance*. 2002. Vol. 26, No. 7. P. 1443-1471.

61. Markowitz H. Portfolio Selection. *The Journal of Finance*. 1952. Vol. 7, No. 1. P. 77-91.

62. Glasserman P. *Monte Carlo Methods in Financial Engineering*. Springer, 2003.

63. McMaster G. S., Wilhelm W. W. Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*. 1997. Vol. 87, No. 4. P. 291-300.

64. Porter J. R., Gawith M. Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy*. 1999. Vol. 10, No. 1. P. 23-36.
65. Sánchez B., Rasmussen A., Porter J. R. Temperatures and the growth and development of maize and rice: a review. *Global Change Biology*. 2014. Vol. 20, No. 2. P. 408-417.
66. Chen T., Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2016. P. 785-794. DOI: 10.1145/2939672.2939785
67. Friedman J. H. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics*. 2001. Vol. 29, No. 5. P. 1189-1232.
68. Hastie T., Tibshirani R., Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer, 2009.
69. Raschka S., Patterson J., Nolet C. Machine Learning in Python: Main Developments and Technology Trends in Data Science, Machine Learning, and Artificial Intelligence. *Information*. 2020. Vol. 11, No. 4. Article 193. DOI: 10.3390/info11040193
70. Pedregosa F., Varoquaux G., Gramfort A. et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011. Vol. 12. P. 2825-2830.
71. Microsoft. .NET documentation. URL: <https://learn.microsoft.com/en-us/dotnet/>
72. Albahari J. *C# 12 in a Nutshell: The Definitive Reference*. O'Reilly Media, 2024.
73. Microsoft. Windows Presentation Foundation (WPF) Documentation. URL: <https://learn.microsoft.com/en-us/dotnet/desktop/wpf/>
74. ScottPlot: Interactive Plotting Library for .NET. URL: <https://scottplot.net/>
75. McKinney W. Data Structures for Statistical Computing in Python. *Proceedings of the 9th Python in Science Conference*. 2010. P. 56-61.
76. Harris C. R., Millman K. J., van der Walt S. J. et al. Array programming with NumPy. *Nature*. 2020. Vol. 585. P. 357-362. DOI: 10.1038/s41586-020-2649-2

77. Kohavi R. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. Proceedings of the 14th International Joint Conference on Artificial Intelligence. 1995. Vol. 2. P. 1137-1143.
78. Fowler M. Patterns of Enterprise Application Architecture. Addison-Wesley, 2002.
79. Martin R. C. Agile Software Development, Principles, Patterns, and Practices. Pearson, 2002.
80. Gamma E., Helm R., Johnson R., Vlissides J. Design Patterns: Elements of Reusable Object-Oriented Software. Addison-Wesley, 1994.
81. Fowler M. Inversion of Control Containers and the Dependency Injection pattern. 2004. URL: <https://martinfowler.com/articles/injection.html>
82. Seemann M., van Deursen S. Dependency Injection Principles, Practices, and Patterns. Manning Publications, 2019.
83. Cleary S. Concurrency in C# Cookbook: Asynchronous, Parallel, and Multithreaded Programming. 2nd ed. O'Reilly Media, 2019.
84. Microsoft. Configuration in .NET. URL: <https://learn.microsoft.com/en-us/dotnet/core/extensions/configuration>
85. Microsoft. Microsoft Edge WebView2 documentation. URL: <https://learn.microsoft.com/en-us/microsoft-edge/webview2/>
86. Breiman L. Random Forests. Machine Learning. 2001. Vol. 45, No. 1. P. 5-32. DOI: 10.1023/A:1010933404324
87. Breiman L. Bagging Predictors. Machine Learning. 1996. Vol. 24, No. 2. P. 123-140.
88. Tibshirani R. Regression Shrinkage and Selection via the Lasso. Journal of the Royal Statistical Society. Series B. 1996. Vol. 58, No. 1. P. 267-288.
89. Hoerl A. E., Kennard R. W. Ridge Regression: Biased Estimation for Nonorthogonal Problems. Technometrics. 1970. Vol. 12, No. 1. P. 55-67.
90. Aggarwal C. C. Outlier Analysis. 2nd ed. Springer, 2017.

91. Little R. J. A., Rubin D. B. Statistical Analysis with Missing Data. 3rd ed. Wiley, 2019.
92. Microsoft. System.Text.Json overview. URL: <https://learn.microsoft.com/en-us/dotnet/standard/serialization/system-text-json/overview>
93. XGBoostSharp: A .NET wrapper for XGBoost. GitHub repository. URL: <https://github.com/PicNet/XGBoost.Net>
94. Microsoft. Native interoperability (.NET documentation). URL: <https://learn.microsoft.com/en-us/dotnet/standard/native-interop/>
95. Agafonkin V. Leaflet – a JavaScript library for interactive maps. URL: <https://leafletjs.com/>
96. Box G. E. P., Muller M. E. A Note on the Generation of Random Normal Deviates. The Annals of Mathematical Statistics. 1958. Vol. 29, No. 2. P. 610-611.
97. Efron B., Tibshirani R. J. An Introduction to the Bootstrap. Chapman and Hall/CRC, 1993.
98. IEEE Std 1012-2016. IEEE Standard for System, Software, and Hardware Verification and Validation. IEEE, 2017.
99. Saltelli A., Ratto M., Andres T. et al. Global Sensitivity Analysis: The Primer. Wiley, 2008.
100. Sujatha V. et al. Assessment of sensor based precision nitrogen management in maize // Maydica. 2024. Vol. 67. No. 16. <https://creajournals.crea.gov.it/index.php/maydica/article/download/2422/2240>
101. Gao, J., Zeng, W., Ren, Z., Ao, C., Lei, G., Gaiser, T., & Srivastava, A. K. (2023). A Fertilization Decision Model for Maize, Rice, and Soybean Based on Machine Learning and Swarm Intelligent Search Algorithms. Agronomy, 13(5), 1400. <https://doi.org/10.3390/agronomy13051400>
102. Ren, Z., Zhang, H., Li, H., Wu, Q., Huang, Y., Ye, Y., & Zhao, Y. (2025). Improving Wheat Yield, Fertilizer Use Efficiency, and Economic Benefits Through

Farmer-Participation Nutrient Management. Sustainability, 17(8), 3481.
<https://doi.org/10.3390/su17083481>

103. Porter, J. R., & Gawith, M. (1999). Temperatures and the growth and development of wheat: a review. *European Journal of Agronomy*, 10(1), 23–36.
[https://doi.org/10.1016/S1161-0301\(98\)00047-1](https://doi.org/10.1016/S1161-0301(98)00047-1)

104. Eyshi Rezaei, E., Webber, H., Gaiser, T., Naab, J., & Ewert, F. (2015). Heat stress in cereals: Mechanisms and modelling. *European Journal of Agronomy*, 64, 98–113. <https://doi.org/10.1016/j.eja.2014.10.003>

105. Jagadish, S. V. K., Craufurd, P. Q., & Wheeler, T. R. (2007). High temperature stress and spikelet fertility in rice (*Oryza sativa* L.). *Journal of Experimental Botany*, 58(7), 1627–1635. <https://doi.org/10.1093/jxb/erm003>

106. Satake, T., & Yoshida, S. (1978). High temperature-induced sterility in indica rices at flowering. *Japanese Journal of Crop Science*, 47(1), 6–17.
<https://doi.org/10.1626/jcs.47.6>

107. McMaster, G. S., & Wilhelm, W. W. (1997). Growing degree-days: one equation, two interpretations. *Agricultural and Forest Meteorology*, 87(4), 291–300. [https://doi.org/10.1016/S0168-1923\(97\)00027-0](https://doi.org/10.1016/S0168-1923(97)00027-0)

ДОДАТКИ

Реалізація сервісу **DataMergerService** для агрегації агростатистичних даних із заздалегідь підготовлених даних **FAOSTAT**

```

public interface IDataMergerService
{
    List<MergedData> MergeFertilizerAndYieldData(string
fertilizerFilePath, string yieldFilePath);
    void SaveMergedData(List<MergedData> data, string outputPath);
}

public class DataMergerService : IDataMergerService
{
    private readonly ICsvService _csvService;

    public DataMergerService(ICsvService csvService)
    {
        _csvService = csvService;
    }

    public List<MergedData> MergeFertilizerAndYieldData(string
fertilizerFilePath, string yieldFilePath)
    {
        var fertilizers =
_csvService.ReadRecords<CsvItemModel>(fertilizerFilePath);
        var yields =
_csvService.ReadRecords<CsvItemModel>(yieldFilePath);

        var groupedFertilizers = fertilizers
            .GroupBy(u => new KeyModel { Year = u.Year, Area = u.Area
        })
            .Select(g => new FertilizerData
            {
                Year = int.Parse(g.Key.Year),
                Country = g.Key.Area,
                N_Used = g.Where(x =>
x.Item.Contains("nitrogen")).Sum(u => double.Parse(u.Value,
CultureInfo.InvariantCulture)),
                P_Used = g.Where(x =>
x.Item.Contains("phosphate")).Sum(u => double.Parse(u.Value,
CultureInfo.InvariantCulture)),
                K_Used = g.Where(x =>
x.Item.Contains("potash")).Sum(u => double.Parse(u.Value,
CultureInfo.InvariantCulture))
            })
            .OrderBy(r => r.Year)
            .ThenBy(r => r.Country)
            .ToList();

        var groupedYields = yields
            .GroupBy(u => new { u.Year, u.Area })
            .Select(g => new YieldData
            {

```

```

        Year = int.Parse(g.Key.Year),
        Country = g.Key.Area,
        YieldGot = g.Sum(u => double.Parse(u.Value,
CultureInfo.InvariantCulture))
    })
    .OrderBy(r => r.Year)
    .ThenBy(r => r.Country)
    .ToList();

    var mergedResult = from fertilizer in groupedFertilizers
                        join yield in groupedYields
                        on new { fertilizer.Year,
fertilizer.Country } equals new { yield.Year, yield.Country }
                        select new MergedData
                        {
                            Year = fertilizer.Year,
                            Country = fertilizer.Country,
                            N_Used = fertilizer.N_Used,
                            P_Used = fertilizer.P_Used,
                            K_Used = fertilizer.K_Used,
                            YieldGot = yield.YieldGot,
                            AvgTempGrowingSeason = 0, // Will be
populated with weather data
                            TotalRainGrowingSeason = 0,
                            GDD = 0,
                            SDD = 0
                        };

    return mergedResult.ToList();
}

public void SaveMergedData(List<MergedData> data, string
outputPath)
{
    var csv = new StringBuilder();

    // Header

    csv.AppendLine("Year,Country,N_Used,P_Used,K_Used,YieldGot,AvgTempGrowing
Season,TotalRainGrowingSeason,GDD,SDD");

    // Data rows
    foreach (var row in data)
    {
        csv.AppendLine($"{row.Year},{row.Country}," +
        $"{row.N_Used.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.P_Used.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.K_Used.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.YieldGot.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.AvgTempGrowingSeason.ToString(CultureInfo.InvariantCulture)}," +

```

```

    $"{row.TotalRainGrowingSeason.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.GDD.ToString(CultureInfo.InvariantCulture)}," +
        $"{row.SDD.ToString(CultureInfo.InvariantCulture)}");
    }
    _csvService.WriteTextToFile(outputPath, csv.ToString());
}
}

```

Додаток Б

Реалізація сервісу WeatherService для збору архівних метеорологічних даних та обчислення сезонних агрокліматичних індикаторів

```

public interface IWeatherService
{
    Task<ClimateData> GetWeather(double lat, double lon, int year);
    Task<List<ClimateData>> GetWeatherForYears(double lat, double lon, int startYear, int endYear);
    Task<Dictionary<string, List<ClimateData>>> GetAllCountriesClimateData();
    Task WriteAllCountriesClimateData(string countriesFilePath, string outputPath);
}

public class WeatherService : IWeatherService
{
    private readonly IHttpService _httpService;
    private readonly ICsvService _csvService;
    private readonly CoreConfiguration _configuration;
    private readonly Dictionary<string, GeoPoint> _agriculturalCentroids;

    public WeatherService(IHttpService httpService, ICsvService csvService, CoreConfiguration configuration)
    {
        _httpService = httpService;
        _csvService = csvService;
        _configuration = configuration;
        _agriculturalCentroids = GetAgriculturalCentroids();
    }

    public async Task<ClimateData> GetWeather(double lat, double lon, int year)
    {
        var weatherData = await GetWeatherForYears(lat, lon, year, year);
        return weatherData.FirstOrDefault() ?? new ClimateData { Year = year };
    }

    public async Task<List<ClimateData>> GetWeatherForYears(double lat, double lon, int startYear, int endYear)

```

```

    {
        string url = PrepareLink(lat, lon, startYear, endYear);
        var response = await _httpClientService.GetStringAsync(url);
        return ParseWeatherResponse(response, startYear, endYear);
    }

    public async Task<Dictionary<string, List<ClimateData>>>
    GetAllCountriesClimateData()
    {
        var result = new Dictionary<string, List<ClimateData>>>();

        foreach (var country in _agriculturalCentroids)
        {
            try
            {
                var climateData = await
                GetWeatherForYears(country.Value.Latitude, country.Value.Longitude, 1990,
                2023);

                result.Add(country.Key, climateData);
            }
            catch (Exception ex)
            {
                Console.WriteLine($"Failed to get weather data for
                {country.Key}: {ex.Message}");
            }
        }

        return result;
    }

    public async Task WriteAllCountriesClimateData(string
    countriesFilePath, string outputPath)
    {
        var countriesCoordinates =
        _csvService.ReadRecords<GeoPointsData>(countriesFilePath);

        var csv = new StringBuilder();

        csv.AppendLine("Country,Year,AvgTempGrowingSeason,TotalRainGrowingSeason,
        GDD,SDD");

        foreach (var country in countriesCoordinates)
        {
            Console.WriteLine($"Fetching {country.Country}...");

            // Простая логика повторы (Retry logic)
            int attempts = 0;
            bool success = false;

            while (!success && attempts < 3)
            {
                try
                {
                    var lat = double.Parse(country.Latitude,
                    CultureInfo.InvariantCulture);

```

```

        var lon = double.Parse(country.Longitude,
CultureInfo.InvariantCulture);
        if
(_agriculturalCentroids.TryGetValue(country.Country, out var
agriculturalCentroid))
        {
            lat = agriculturalCentroid.Latitude;
            lon = agriculturalCentroid.Longitude;
        }
        var itemWeatherData = await
GetWeatherForYears(lat, lon, 1990, 2023);

        foreach (var itemWeather in itemWeatherData)
        {

csv.AppendLine($"{country.Country},{itemWeather.Year},{itemWeather.AvgTem
pGrowingSeason.ToString(CultureInfo.InvariantCulture)},{itemWeather.Total
RainGrowingSeason.ToString(CultureInfo.InvariantCulture)},{itemWeather.GD
D},{itemWeather.SDD}");
        }

        success = true;
        Console.WriteLine(" Done.");
    }
    catch (HttpRequestException ex)
    {
        Console.WriteLine(ex.Message);
        attempts++;
        Console.WriteLine($"!!! Отримано 429 (Too Many
Requests). Пауза 60 секунд...");
        // Backup current progress

_csvService.WriteTextToFile($"{_configuration.BackupDirectory}\\weather_d
ata{Guid.NewGuid()}.txt", csv.ToString());
        await Task.Delay(60000);
    }
    catch (Exception ex)
    {
        attempts++;
        Console.WriteLine($" Error (Attempt {attempts}):
{ex.Message}");
        // Backup current progress

_csvService.WriteTextToFile($"{_configuration.BackupDirectory}\\weather_d
ata{Guid.NewGuid()}.txt", csv.ToString());
        // Якщо помилка - чекаємо довше (експоненціальна
затримка: 11с, 22с, 33с)
        Thread.Sleep(11000 * attempts);
    }
}

await Task.Delay(10000);

_csvService.WriteTextToFile(outputPath, csv.ToString());
}

```

```

        public void ParseWeatherDataFromFiles(StringBuilder csv)
        {
            var files =
Directory.GetFiles(_configuration.CountriesFilesDirectory);
            foreach (var file in files)
            {
                var countryName = Path.GetFileNameWithoutExtension(file);
                var jsonData = File.ReadAllText(file);
                var itemWeatherData = ParseWeatherResponse(jsonData,
1990, 2023);
                foreach (var itemWeather in itemWeatherData)
                {
                    csv.AppendLine($"{countryName},{itemWeather.Year},{itemWeather.AvgTempGro
wingSeason.ToString(CultureInfo.InvariantCulture)},{itemWeather.TotalRain
GrowingSeason.ToString(CultureInfo.InvariantCulture)},{itemWeather.GDD},{
itemWeather.SDD}");
                }
            }
        }

        private List<ClimateData> ParseWeatherResponse(string response,
int startYear, int endYear)
        {
            const double MIN_TEMP_MEAN = 5.0;

            using var doc = JsonDocument.Parse(response);
            doc.RootElement.GetProperty("latitude").TryGetDouble(out var
lat);

            // Dynamic stress threshold based on latitude
            double stressThreshold = CalculateStressThreshold(lat);

            var daily = doc.RootElement.GetProperty("daily");

            var dates =
daily.GetProperty("time").EnumerateArray().Select(x =>
x.GetDateTime()).ToList();
            var tempsMean =
daily.GetProperty("temperature_2m_mean").EnumerateArray()
                .Select(x => x.ValueKind ==
JsonValueKind.Null ? 0 : x.GetDouble()).ToList();
            var tempsMax =
daily.GetProperty("temperature_2m_max").EnumerateArray()
                .Select(x => x.ValueKind ==
JsonValueKind.Null ? 0 : x.GetDouble()).ToList();

            var rains = daily.GetProperty("rain_sum").EnumerateArray()
                .Select(x => x.ValueKind ==
JsonValueKind.Null ? 0 : x.GetDouble()).ToList();

            var result = new List<ClimateData>();

            for (int y = startYear; y <= endYear; y++)
            {

```

```

var yearIndices = dates.Select((date, index) => new {
date, index })
                                .Where(x => x.date.Year == y)
                                .ToList();
var activeGrowingDays = yearIndices
    .Where(x => tempsMean[x.index] >= MIN_TEMP_MEAN)
    .ToList();

if (activeGrowingDays.Any())
{
    double totalRain = activeGrowingDays.Sum(x =>
rains[x.index]);
    double avgTempSeason = activeGrowingDays.Average(x =>
tempsMean[x.index]);

    double gddSum = 0;
    double sddSum = 0;

    foreach (var day in activeGrowingDays)
    {
        double tMean = tempsMean[day.index];
        double tMax = tempsMax[day.index];
        if (tMean > MIN_TEMP_MEAN)
        {
            gddSum += (tMean - MIN_TEMP_MEAN);
        }
        if (tMax > stressThreshold)
        {
            sddSum += (tMax - stressThreshold);
        }
    }

    result.Add(new ClimateData
    {
        Year = y,
        AvgTempGrowingSeason = Math.Round(avgTempSeason,
2),
        TotalRainGrowingSeason = Math.Round(totalRain,
2),
        GrowingDaysCount = activeGrowingDays.Count,
        HeatStressDaysCount = activeGrowingDays.Count(x
=> tempsMax[x.index] >= stressThreshold),
        GDD = Math.Round(gddSum, 2),
        SDD = Math.Round(sddSum, 2)
    });
    }
}
return result;
}

private double CalculateStressThreshold(double latitude)
{
    double absLat = Math.Abs(latitude);

    if (absLat < 20)
    {

```

```

        return 34.0;
    }
    else if (absLat >= 20 && absLat < 50)
    {
        return 34.0 - 4 * ((absLat - 20) / 30);
    }
    else
    {
        return 30.0;
    }
}

private string PrepareLink(double lat, double lon, int startYear,
int endYear)
{
    string startDate = $"{startYear}-01-01";
    string endDate = $"{endYear}-12-31";
    string url = $"https://archive-api.open-
meteo.com/v1/archive?latitude={lat.ToString(CultureInfo.InvariantCulture)}
&longitude={lon.ToString(CultureInfo.InvariantCulture)}&start_date={star
tDate}&end_date={endDate}&daily=temperature_2m_mean,temperature_2m_max,ra
in_sum&timezone=auto";
    return url;
}

private Dictionary<string, GeoPoint> GetAgriculturalCentroids()
{
    return new Dictionary<string, GeoPoint>
    {
        // ПІВНІЧНА АМЕРИКА
        { "Canada", new GeoPoint { Latitude = 52.1, Longitude = -
106.6 } },
        { "United States of America", new GeoPoint { Latitude =
42.0, Longitude = -93.6 } },
        // ПІВДЕННА АМЕРИКА
        { "Brazil", new GeoPoint { Latitude = -13.0, Longitude =
-56.0 } },
        { "Argentina", new GeoPoint { Latitude = -34.6, Longitude
= -60.9 } },
        { "Chile", new GeoPoint { Latitude = -35.4, Longitude = -
71.6 } },
        { "Peru", new GeoPoint { Latitude = -8.1, Longitude = -
79.0 } },
        // ЄВРОПА (Північ та Схід)
        { "Russian Federation", new GeoPoint { Latitude = 51.6,
Longitude = 39.2 } },
        { "Ukraine", new GeoPoint { Latitude = 49.4, Longitude =
32.0 } },
        { "Sweden", new GeoPoint { Latitude = 58.4, Longitude =
13.8 } },
        { "Finland", new GeoPoint { Latitude = 61.0, Longitude =
24.0 } },
        { "Norway", new GeoPoint { Latitude = 59.9, Longitude =
10.7 } },
    }
}

```

```

// АФРИКА
{ "Algeria", new GeoPoint { Latitude = 35.3, Longitude =
1.3 } },
{ "Egypt", new GeoPoint { Latitude = 30.9, Longitude =
30.6 } },
{ "Libya", new GeoPoint { Latitude = 32.8, Longitude =
13.1 } },
{ "Morocco", new GeoPoint { Latitude = 33.8, Longitude =
-5.5 } },
{ "Tunisia", new GeoPoint { Latitude = 36.8, Longitude =
9.5 } },
{ "Sudan", new GeoPoint { Latitude = 14.4, Longitude =
33.5 } },
{ "Sudan (former)", new GeoPoint { Latitude = 14.4,
Longitude = 33.5 } },

// БЛИЗЬКИЙ СХІД
{ "Saudi Arabia", new GeoPoint { Latitude = 26.3,
Longitude = 43.9 } },
{ "Iraq", new GeoPoint { Latitude = 32.5, Longitude =
45.8 } },
{ "Syrian Arab Republic", new GeoPoint { Latitude = 36.2,
Longitude = 37.1 } },
{ "Iran", new GeoPoint { Latitude = 34.3, Longitude =
47.0 } },

// ЦЕНТРАЛЬНА АЗІЯ
{ "Kazakhstan", new GeoPoint { Latitude = 53.2, Longitude
= 63.6 } },
{ "Uzbekistan", new GeoPoint { Latitude = 40.1, Longitude
= 67.8 } },
{ "Turkmenistan", new GeoPoint { Latitude = 37.5,
Longitude = 61.8 } },

// АЗІЯ / ОКЕАНІЯ
{ "China", new GeoPoint { Latitude = 34.7, Longitude =
113.6 } },
{ "Australia", new GeoPoint { Latitude = -34.2, Longitude
= 146.0 } },
{ "India", new GeoPoint { Latitude = 26.8, Longitude =
80.9 } },
{ "Pakistan", new GeoPoint { Latitude = 30.8, Longitude =
72.4 } },
{ "Turkey", new GeoPoint { Latitude = 39.0, Longitude =
32.8 } },

// ЄВРОПА (Доповнення)
{ "France", new GeoPoint { Latitude = 48.4, Longitude =
1.5 } },
{ "Romania", new GeoPoint { Latitude = 44.5, Longitude =
27.0 } },

// АФРИКА (Доповнення)
{ "Ethiopia", new GeoPoint { Latitude = 8.9, Longitude =
38.7 } },

```

```

        { "South Africa", new GeoPoint { Latitude = -28.4,
Longitude = 26.5 } },

        // IHIII
        { "Mexico", new GeoPoint { Latitude = 20.5, Longitude = -
101.5 } },
        { "Indonesia", new GeoPoint { Latitude = -7.5, Longitude
= 111.0 } },
        { "Mauritania", new GeoPoint { Latitude = 16.5, Longitude
= -15.8 } },
        { "Niger", new GeoPoint { Latitude = 13.5, Longitude =
7.1 } },
        { "Chad", new GeoPoint { Latitude = 9.1, Longitude = 18.4
} },
        { "Mali", new GeoPoint { Latitude = 13.4, Longitude = -
6.2 } },
        { "Senegal", new GeoPoint { Latitude = 14.1, Longitude =
-16.0 } },
        { "Burkina Faso", new GeoPoint { Latitude = 12.3,
Longitude = -1.5 } },

        { "Nigeria", new GeoPoint { Latitude = 11.0, Longitude =
7.7 } },

        // --- СХІДНА АФРИКА ---
        { "Kenya", new GeoPoint { Latitude = -0.3, Longitude =
36.0 } },
        { "Tanzania", new GeoPoint { Latitude = -6.8, Longitude =
37.6 } },
        { "Mongolia", new GeoPoint { Latitude = 49.0, Longitude =
105.0 } },
        { "Japan", new GeoPoint { Latitude = 37.9, Longitude =
139.0 } },
        { "Vietnam", new GeoPoint { Latitude = 10.5, Longitude =
105.6 } },

        // --- ЛАТИНСЬКА АМЕРИКА ---
        { "Colombia", new GeoPoint { Latitude = 4.0, Longitude =
-75.0 } },
        { "Venezuela", new GeoPoint { Latitude = 9.0, Longitude =
-67.5 } }
    };
}

```

Додаток В

Програмна реалізація тренування та серіалізації прогновної моделі, порівняння алгоритмів машинного навчання та абляційного аналізу

```

import pandas as pd
import numpy as np
from sklearn.model_selection import train_test_split, KFold

```

```

from sklearn.metrics import mean_squared_error, r2_score,
mean_absolute_error
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
from xgboost import XGBRegressor
import json

def add_ctl_from_json(df_train, df_test,
json_path='country_tech_levels.json'):
    with open(json_path, 'r', encoding='utf-8') as f:
        ctl_map = json.load(f)
    global_fallback = sum(ctl_map.values()) / len(ctl_map)
    df_train = df_train.copy()
    df_test = df_test.copy()
    df_train['CountryTechLevel'] = (
        df_train['Country'].map(ctl_map).fillna(global_fallback)
    )
    df_test['CountryTechLevel'] = (
        df_test['Country'].map(ctl_map).fillna(global_fallback)
    )
    return df_train, df_test

def prepare_data_leak_free(df, include_ctl=True, include_year=False,
random_state=42):
    q_low = df["YieldGot"].quantile(0.01)
    q_hi = df["YieldGot"].quantile(0.99)
    df = df[(df["YieldGot"] < q_hi) & (df["YieldGot"] > q_low)].copy()
    base_features = ['N_Used', 'P_Used', 'K_Used',
                    'TotalRainGrowingSeason', 'GDD', 'SDD']
    needed_cols = base_features + ['YieldGot', 'Country']
    if include_year:
        needed_cols.append('Year')
    df = df.dropna(subset=needed_cols)
    df_train, df_test = train_test_split(df, test_size=0.2,
random_state=random_state)
    if include_ctl:
        df_train, df_test = add_ctl_from_json(df_train, df_test)
    features = list(base_features)
    if include_ctl:
        features.append('CountryTechLevel')
    if include_year:
        features.append('Year')
    X_train = df_train[features]
    X_test = df_test[features]
    y_train_real = df_train['YieldGot'].values
    y_test_real = df_test['YieldGot'].values
    y_train_log = np.log1p(y_train_real)
    y_test_log = np.log1p(y_test_real)
    return X_train, X_test, y_train_log, y_test_log, y_train_real,
y_test_real, features

def evaluate_model(model, X_train, X_test, y_train_log, y_test_log,
                    y_train_real, y_test_real, config_name=""):
    model.fit(X_train, y_train_log)
    y_pred_train_log = model.predict(X_train)
    y_pred_test_log = model.predict(X_test)

```

```

y_pred_train_real = np.expml(y_pred_train_log)
y_pred_test_real = np.expml(y_pred_test_log)
y_pred_train_real = np.clip(y_pred_train_real, 0, None)
y_pred_test_real = np.clip(y_pred_test_real, 0, None)
metrics = {
    "config": config_name,
    "Train RMSE": np.sqrt(mean_squared_error(y_train_real,
y_pred_train_real)),
    "Test RMSE": np.sqrt(mean_squared_error(y_test_real,
y_pred_test_real)),
    "Train MAE": mean_absolute_error(y_train_real,
y_pred_train_real),
    "Test MAE": mean_absolute_error(y_test_real, y_pred_test_real),
    "Train R2": r2_score(y_train_real, y_pred_train_real),
    "Test R2": r2_score(y_test_real, y_pred_test_real),
}
residuals = y_test_log - y_pred_test_log
metrics["Sigma^2"] = float(np.var(residuals))
metrics["0.5*Sigma^2"] = 0.5 * metrics["Sigma^2"]
return metrics, model

def honest_cv(df, model_factory, include_ctl=True, include_year=False,
              n_splits=5, random_state=42):
    q_low = df["YieldGot"].quantile(0.01)
    q_hi = df["YieldGot"].quantile(0.99)
    df = df[(df["YieldGot"] < q_hi) & (df["YieldGot"] > q_low)].copy()
    base_features = ['N_Used', 'P_Used', 'K_Used',
                    'TotalRainGrowingSeason', 'GDD', 'SDD']
    needed_cols = base_features + ['YieldGot', 'Country']
    if include_year:
        needed_cols.append('Year')
    df = df.dropna(subset=needed_cols).reset_index(drop=True)
    features = list(base_features)
    if include_ctl:
        features.append('CountryTechLevel')
    if include_year:
        features.append('Year')
    kf = KFold(n_splits=n_splits, shuffle=True,
random_state=random_state)
    r2_scores, rmse_scores, mae_scores = [], [], []
    for train_idx, val_idx in kf.split(df):
        df_tr, df_val = df.iloc[train_idx], df.iloc[val_idx]
        if include_ctl:
            df_tr, df_val = add_ctl_from_json(df_tr, df_val)
        X_tr = df_tr[features]
        X_val = df_val[features]
        y_tr_log = np.log1p(df_tr['YieldGot'].values)
        y_val_real = df_val['YieldGot'].values
        model = model_factory()
        model.fit(X_tr, y_tr_log)
        y_pred_real = np.clip(np.expml(model.predict(X_val)), 0, None)
        r2_scores.append(r2_score(y_val_real, y_pred_real))
        rmse_scores.append(np.sqrt(mean_squared_error(y_val_real,
y_pred_real)))
        mae_scores.append(mean_absolute_error(y_val_real, y_pred_real))

```

```

return {
    "CV R2 mean":    np.mean(r2_scores),
    "CV R2 std":     np.std(r2_scores),
    "CV RMSE mean":  np.mean(rmse_scores),
    "CV RMSE std":   np.std(rmse_scores),
    "CV MAE mean":   np.mean(mae_scores),
    "CV MAE std":    np.std(mae_scores),
    "R2 per fold":   r2_scores,
}

# Фабрики моделей для CV
def mk_lr():    return LinearRegression()
def mk_rf():    return RandomForestRegressor(n_estimators=200,
max_depth=None,
                                           n_jobs=-1, random_state=42)
def mk_xgb():   return XGBRegressor(n_estimators=500, learning_rate=0.05,
max_depth=6, n_jobs=-1,
random_state=42)
def run_model_comparison(file_path):
    df_raw = pd.read_csv(file_path)
    X_train, X_test, y_tr_log, y_te_log, y_tr_real, y_te_real, feats = \
        prepare_data_leak_free(df_raw, include_ctl=True,
include_year=False)
    print(f"\n{'=' * 70}")
    print("ПОРІВНЯННЯ МОДЕЛЕЙ (CTL leak-free, без Year) – для СТАТТИ")
    print(f"{'=' * 70}")
    print(f"Фічі: {feats}")
    print(f"Train size: {len(X_train)} | Test size: {len(X_test)}")
    model_specs = [
        ("Linear Regression", mk_lr, None),
        ("Random Forest",     mk_rf, None),
        ("XGBoost",           mk_xgb, "xgboost_model_leakfree.bin"),
    ]
    results = []
    for name, factory, save_path in model_specs:
        print(f"\n--- {name} ---")
        model = factory()
        metrics, trained = evaluate_model(
            model, X_train, X_test, y_tr_log, y_te_log,
            y_tr_real, y_te_real, config_name=name
        )
        print(f"  Train RMSE: {metrics['Train RMSE']:7.2f} | Test RMSE:
{metrics['Test RMSE']:7.2f}")
        print(f"  Train MAE: {metrics['Train MAE']:7.2f} | Test
MAE: {metrics['Test MAE']:7.2f}")
        print(f"  Train R²: {metrics['Train R2']:.4f} | Test
R²: {metrics['Test R2']:.4f}")
        cv = honest_cv(df_raw, factory, include_ctl=True,
include_year=False)
        print(f"  CV R²: {cv['CV R2 mean']:.4f} ± {cv['CV R2
std']:.4f}")
        print(f"  CV RMSE: {cv['CV RMSE mean']:.2f} ± {cv['CV RMSE
std']:.2f}")
        print(f"  CV MAE: {cv['CV MAE mean']:.2f} ± {cv['CV MAE
std']:.2f}")

```

```

        metrics.update({
            "CV R2 mean":    cv['CV R2 mean'],
            "CV R2 std":     cv['CV R2 std'],
            "CV RMSE mean":  cv['CV RMSE mean'],
            "CV MAE mean":   cv['CV MAE mean'],
        })
        results.append(metrics)

        if save_path and isinstance(trained, XGBRegressor):
            trained.get_booster().save_model(save_path)
    print(f"\n{'=' * 70}")
    print("ПІДСУМКОВА ТАБЛИЦЯ ПОРІВНЯННЯ МОДЕЛЕЙ")
    print(f"{'=' * 70}")
    df_out = pd.DataFrame(results)
    print(df_out[['config', 'Test R2', 'Test RMSE', 'Test MAE',
                  'CV R2 mean', 'CV RMSE mean', 'CV MAE
mean']].to_string(index=False))
    df_out.to_csv("model_comparison.csv", index=False)
    return df_out

def run_ablation_study(file_path):
    df_raw = pd.read_csv(file_path)
    configs = [
        ("baseline (без CTL, без Year)", False, False),
        ("+ CTL (leak-free)", True, False),
        ("+ Year (без CTL)", False, True),
        ("+ CTL + Year (leak-free)", True, True),
    ]
    results = []
    for name, use_ctl, use_year in configs:
        print(f"\n{'=' * 70}")
        print(f"КОНФІГУРАЦІЯ: {name}")
        print(f"{'=' * 70}")
        X_train, X_test, y_tr_log, y_te_log, y_tr_real, y_te_real, feats
= \
            prepare_data_leak_free(df_raw, include_ctl=use_ctl,
include_year=use_year)
        print(f"Фічі: {feats} | Train: {len(X_train)}, Test:
{len(X_test)}")
        metrics, _ = evaluate_model(
            mk_xgb(), X_train, X_test, y_tr_log, y_te_log,
            y_tr_real, y_te_real, config_name=name
        )
        print(f"  Train RMSE: {metrics['Train RMSE']:7.2f} | Test RMSE:
{metrics['Test RMSE']:7.2f}")
        print(f"  Train MAE: {metrics['Train MAE']:7.2f} | Test
MAE: {metrics['Test MAE']:7.2f}")
        print(f"  Train R²: {metrics['Train R2']:.4f} | Test
R²: {metrics['Test R2']:.4f}")
        cv = honest_cv(df_raw, mk_xgb, include_ctl=use_ctl,
include_year=use_year)
        print(f"  CV R²: {cv['CV R2 mean']:.4f} ± {cv['CV R2
std']:.4f}")
        print(f"  CV RMSE: {cv['CV RMSE mean']:.2f} ± {cv['CV RMSE
std']:.2f}")

```

```

        print(f"   CV MAE:   {cv['CV MAE mean']:.2f} ± {cv['CV MAE
std']:.2f}")
        metrics.update({
            "CV R2 mean":   cv['CV R2 mean'],
            "CV R2 std":    cv['CV R2 std'],
            "CV RMSE mean": cv['CV RMSE mean'],
            "CV MAE mean":  cv['CV MAE mean'],
        })
        results.append(metrics)
    print(f"\n{'=' * 70}")
    print("ПІДСУМКОВА ТАБЛИЦЯ ABLATION")
    print('=' * 70)
    df_out = pd.DataFrame(results)
    print(df_out[['config', 'Test R2', 'Test RMSE', 'Test MAE',
                  'CV R2 mean', 'CV RMSE mean', 'CV MAE
mean']].to_string(index=False))
    df_out.to_csv("ablation_results.csv", index=False)
    return df_out

if __name__ == "__main__":
    _ = run_model_comparison('FULL_DATA.csv')

    _ = run_ablation_study('FULL_DATA.csv')

```

Додаток Г

Реалізація сервісу `ModelPredictorService` для нативного інференсу прогновної моделі `XGBoost` у керованому коді клієнтського застосунку

```

public interface IModelPredictorService
{
    float[] PredictFromMergedData(List<MergedData> data);
}

public class ModelPredictorService : IModelPredictorService
{
    // from learning
    const float _biasCorrection = 0.018769869132406174f;
    private readonly string _modelPath;

    public ModelPredictorService(CoreConfiguration configuration)
    {
        _modelPath = configuration.ModelPath;

        if (!File.Exists(_modelPath))
            throw new FileNotFoundException($"Model file not found:
{_modelPath}");
    }

    public float[] PredictFromMergedData(List<MergedData> input)
    {
        float[][] batchInputData = new float[input.Count][];
        for (int i = 0; i < input.Count; i++)
        {

```

```

        var inputItem = input[i];
        batchInputData[i] = [ (float)inputItem.N_Used,
(float)inputItem.P_Used, (float)inputItem.K_Used,
(float)inputItem.TotalRainGrowingSeason, (float)inputItem.GDD,
(float)inputItem.SDD, inputItem.CountryLevel ];
    }

    var model = XGBRegressor.LoadFromFile(_modelPath);
    var expRes = model.Predict(batchInputData);
    var results = new float[expRes.Length];
    foreach (var result in expRes)
        results[Array.IndexOf(expRes, result)] =
(float)Math.Pow(Math.E, result + _biasCorrection);

    return results;
}
}

```

Додаток Д

Реалізація сервісу `ScenarioOptimizer`, що здійснює оптимізацію за прибутковістю з урахуванням ризику та сценарний аналіз

```

public interface IScenarioOptimizer
{
    Task<ScenarioOptimizationResult> RunAsync(ScenarioInput input);
}

public class ScenarioOptimizer : IScenarioOptimizer
{
    private readonly IWeatherService _weatherService;
    private readonly IFertilizerService _fertilizerService;
    private readonly IModelPredictorService _modelPredictorService;
    private readonly ICountryService _countryService;

    public ScenarioOptimizer(
        IWeatherService weatherService,
        IFertilizerService fertilizerService,
        IModelPredictorService modelPredictorService,
        ICountryService countryService)
    {
        _weatherService = weatherService;
        _fertilizerService = fertilizerService;
        _modelPredictorService = modelPredictorService;
        _countryService = countryService;
    }

    public async Task<ScenarioOptimizationResult> RunAsync(ScenarioInput
input)
    {
        int endYear = DateTime.Now.Year - 1;
        int startYear = endYear - input.HistoricalYears + 1;

        List<ClimateData> historicalWeather;
    }
}

```

```

try
{
    historicalWeather = await _weatherService.GetWeatherForYears(
        input.Latitude, input.Longitude, startYear, endYear);
}
catch (Exception ex)
{
    throw new InvalidOperationException(
        $"Не вдалося отримати погодні дані для
({input.Latitude:F4}, {input.Longitude:F4}): {ex.Message}", ex);
}

if (historicalWeather.Count == 0)
    throw new InvalidOperationException(
        $"Погодні дані за {startYear}-{endYear} недоступні для
вказаної локації.");

string? warning = null;
if (historicalWeather.Count < 5)
    warning = $"Попередження: отримано лише
{historicalWeather.Count} річних реалізацій. " +
        "Для адекватності аналізу рекомендується мінімум 5
років.";
var rng = new Random(42);

double[] gddSeries = historicalWeather.Select(c =>
(double)c.GDD).ToArray();
double[] sddSeries = historicalWeather.Select(c =>
(double)c.SDD).ToArray();
double[] rainSeries = historicalWeather.Select(c =>
(double)c.TotalRainGrowingSeason).ToArray();

double meanGDD = gddSeries.Average();
double meanSDD = sddSeries.Average();
double meanRain = rainSeries.Average();

double stdGDD = StandardDeviation(gddSeries);
double stdSDD = StandardDeviation(sddSeries);
double stdRain = StandardDeviation(rainSeries);

float baseAvgTemp = (float)historicalWeather.Average(c =>
c.AvgTempGrowingSeason);
string baseCountry = historicalWeather[0].Country;
var scenarios = new ClimateData[input.ScenarioCount];
for (int i = 0; i < input.ScenarioCount; i++)
{
    scenarios[i] = new ClimateData
    {
        GDD = (float)Math.Max(0, SampleNormal(rng, meanGDD,
stdGDD)),
        SDD = (float)Math.Max(0, SampleNormal(rng, meanSDD,
stdSDD)),
        TotalRainGrowingSeason = (float)Math.Max(0,
SampleNormal(rng, meanRain, stdRain)),
        AvgTempGrowingSeason = baseAvgTemp,
        Year = endYear,
    }
}

```

```

        Country = baseCountry
    };
}
float techLevel = !string.IsNullOrEmpty(input.CountryName)
    ? _countryService.GetTechLevelByName(input.CountryName)
    : await _countryService.GetTechLevelAsync(input.Latitude,
input.Longitude);
var fertilizer =
_fertilizerService.GetFertilizer(input.FertilizerId)
    ?? throw new InvalidOperationException($"Добриво з ID
{input.FertilizerId} не знайдено.");

int dosesCount = input.DosesKgPerHa.Length;
var doseNPKs = new NPKCalculationResult[dosesCount];
var batchInputs = new List<MergedData>(dosesCount *
input.ScenarioCount);

for (int d = 0; d < dosesCount; d++)
{
    doseNPKs[d] =
_fertilizerService.CalculateNPK(input.FertilizerId,
input.DosesKgPerHa[d]);

    for (int s = 0; s < input.ScenarioCount; s++)
    {
        var climate = scenarios[s];
        batchInputs.Add(new MergedData
        {
            N_Used = doseNPKs[d].NitrogenKg,
            P_Used = doseNPKs[d].PhosphorusKg,
            K_Used = doseNPKs[d].PotassiumKg,
            TotalRainGrowingSeason =
climate.TotalRainGrowingSeason,
            GDD = climate.GDD,
            SDD = climate.SDD,
            CountryLevel = techLevel,
            Year = climate.Year,
            Country = climate.Country,
            AvgTempGrowingSeason = climate.AvgTempGrowingSeason
        });
    }
}
var allYields =
_modelPredictorService.PredictFromMergedData(batchInputs);
var doseMetricsList = new List<DoseMetrics>(dosesCount);

for (int d = 0; d < dosesCount; d++)
{
    int offset = d * input.ScenarioCount;
    var npk = doseNPKs[d];
    double fertCost = input.DosesKgPerHa[d] *
fertilizer.PricePerKg;

    var profits = new double[input.ScenarioCount];
    for (int s = 0; s < input.ScenarioCount; s++)

```

```

        profits[s] = input.CulturePrice * allYields[offset + s] -
fertCost - input.FixedCosts;

        double mean = profits.Average();
        double variance = profits.Select(p => Math.Pow(p - mean,
2)).Average();
        double std = Math.Sqrt(variance);
        int cvarCount = Math.Max(1, (int)(input.ScenarioCount *
0.1));
        double cvar10 = profits.OrderBy(p =>
p).Take(cvarCount).Average();

        double mvScore = mean - input.Lambda * std;

        doseMetricsList.Add(new DoseMetrics(
            Dose: input.DosesKgPerHa[d],
            MeanProfit: mean,
            StdProfit: std,
            Variance: variance,
            CVaR10: cvar10,
            MeanVarianceScore: mvScore
        ));
    }
    var optMV = doseMetricsList.MaxBy(m => m.MeanVarianceScore!);
    var optCVaR = doseMetricsList.MaxBy(m => m.CVaR10!);

    return new ScenarioOptimizationResult(
        DoseMetricsList: doseMetricsList,
        OptimalDoseMV: optMV.Dose,
        OptimalDoseCVaR: optCVaR.Dose,
        ScenariosUsed: input.ScenarioCount,
        Warning: warning
    );
}

private static double StandardDeviation(double[] values)
{
    if (values.Length < 2)
        return 0;

    double mean = values.Average();
    double sumSquaredDiff = values.Sum(v => (v - mean) * (v - mean));
    return Math.Sqrt(sumSquaredDiff / (values.Length - 1));
}

private static double SampleNormal(Random rng, double mean, double
std)
{
    double u1 = 1.0 - rng.NextDouble(); // (0, 1]
    double u2 = 1.0 - rng.NextDouble();
    double standardNormal = Math.Sqrt(-2.0 * Math.Log(u1)) *
Math.Sin(2.0 * Math.PI * u2);
    return mean + std * standardNormal;
}
}

```

Програмна реалізація обчислення контекстної ознаки CountryTechLevel (клас CtlCalculator)

```

public interface ICtlCalculator
{
    List<CountryCropCoefficientDataItem> Calculate(string
dataFilePath);
}

public class CtlCalculator: ICtlCalculator
{
    private readonly ICsvService _csvService;
    public CtlCalculator(ICsvService csvService)
    {
        _csvService = csvService;
    }
    public List<CountryCropCoefficientDataItem> Calculate(string
dataFilePath)
    {
        var records =
_csvService.ReadRecords<CropData>(dataFilePath);
        records = records.Where(r => r.Value.HasValue).ToList();
        var globalCropAverages = records
            .GroupBy(r => r.Item)
            .ToDictionary(
                g => g.Key,
                g => g.Average(r => r.Value!.Value));
        var countryAnalysis = records
            .GroupBy(r => r.Area)
            .Select(countryGroup =>
            {
                var countryName = countryGroup.Key;
                var cropCoefficients = countryGroup
                    .GroupBy(r => r.Item)
                    .Select(cropGroup =>
                    {
                        var cropName = cropGroup.Key;
                        double countryCropAverage =
cropGroup.Average(r => r.Value!.Value);
                        double globalAverage =
globalCropAverages[cropName];
                        double coefficient = globalAverage > 0 ?
countryCropAverage / globalAverage : 0;
                        return new CropCoefficient(cropName,
coefficient);
                    })
                    .ToList();
                double finalCountryCoefficient =
cropCoefficients.Average(c => c.Coeff);
            })
    }
}

```

```
                return new
CountryCropCoefficientDataItem(countryName, cropCoefficients,
finalCountryCoefficient);
            }).ToList();
        return countryAnalysis;
    }
}
```